

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 1 月 1 7 日
Date of Application:

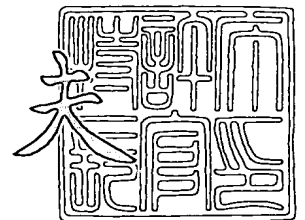
出 願 番 号 特 願 2 0 0 3 - 3 8 6 2 8 7
Application Number:
[ST. 10/C] : [J P 2 0 0 3 - 3 8 6 2 8 7]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 3 年 1 2 月 2 5 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康



【書類名】 特許願
【整理番号】 NT03P0744
【提出日】 平成15年11月17日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 13/10
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 田中 勝也
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 白銀 哲也
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 100068504
 【弁理士】
 【氏名又は名称】 小川 勝男
 【電話番号】 03-3537-1621
【選任した代理人】
 【識別番号】 100086656
 【弁理士】
 【氏名又は名称】 田中 恭助
 【電話番号】 03-3537-1621
【選任した代理人】
 【識別番号】 100094352
 【弁理士】
 【氏名又は名称】 佐々木 孝
 【電話番号】 03-3537-1621
【手数料の表示】
 【予納台帳番号】 081423
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

チャネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなるディスク装置において、

前記ディスクアダプタと前記ディスクアレイを、スイッチを介して接続し、

前記ディスクアダプタと前記ディスクドライブ間を転送されるエクステンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定することを特徴とするディスク装置。

【請求項 2】

請求項 1 記載のディスク装置において、コマンドの種類がデータの読み出しコマンドか、あるいはデータの書き込みコマンドかによって、前記フレームの送信先ドライブポートを決定することを特徴とするディスク装置。

【請求項 3】

請求項 2 記載のディスク装置において、データの読み出しに関する上記エクステンジと、データの書き込みに関する上記エクステンジが、平行して進行することを特徴とするディスク装置。

【請求項 4】

チャネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなるディスク装置において、

前記ディスクアダプタと前記ディスクアレイを、スイッチを介して接続し、

前記ディスクアダプタと前記ディスクドライブ間のエクステンジに含まれるコマンドの種類によって、前記スイッチと前記ディスクドライブ間において転送されるフレームが通るパスを決定することを特徴とするディスク装置。

【請求項 5】

請求項 4 記載のディスク装置において、コマンドの種類がデータの読み出しコマンドか、あるいはデータの書き込みコマンドかによって、前記スイッチと前記ディスクドライブ間において前記フレームが通るパスを決定することを特徴とするディスク装置。

【請求項 6】

チャネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなるディスク装置において、

前記ディスクアダプタと前記ディスクアレイを、スイッチを介して接続し、

前記ディスクアダプタと前記ディスクドライブ間のエクステンジに含まれるコマンドの種類によって、前記ディスクアダプタから前記ディスクドライブへ送るフレーム内の送信先情報を前記ディスクアダプタにおいて決定し、

前記スイッチは、前記ディスクアダプタが接続されたポートと前記ディスクアレイを構成するディスクドライブが接続された各ポートとのポート間接続の切り換えを、入力されたフレーム毎に、該フレーム内の送信先情報に従って行うことを特徴とするディスク装置。

【請求項 7】

チャネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなり、

前記ディスクアダプタと前記ディスクアレイとが、スイッチを介して接続されたディスク装置の制御方法において、

前記スイッチが、前記ディスクアダプタが接続されたポートと前記ディスクアレイを構成するディスクドライブが接続された各ポートとのポート間接続を、入力されたフレーム毎に、前記ディスクアダプタと前記ディスクドライブ間のエクステンジに含まれるコマンドの種類とフレーム内の送信先情報に従って、切り換えることを特徴とするディスク装置

の制御方法。

【請求項 8】

前記スイッチが、前記ディスクアダプタから前記ディスクドライブへ転送されるフレームについては該フレーム内の送信先情報と誤り検出符号を変更し、前記ディスクドライブから前記ディスクアダプタへ転送されるフレームについては該フレーム内の送信元情報と誤り検出符号を変更することを特徴とする請求項 7 記載のディスク装置の制御方法。

【請求項 9】

チャンネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなり、

前記ディスクアダプタと前記ディスクドライブの第 1 のドライブポート群が、第 1 のスイッチを介して接続され、前記ディスクアダプタと前記ディスクドライブの第 2 のドライブポート群が、第 2 のスイッチを介して接続され、さらに第 1 のスイッチと第 2 のスイッチとが接続されたディスク装置の制御方法において、

前記ディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定することを特徴とするディスク装置の制御方法。

【請求項 10】

チャンネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなり、

第 1 のディスクアダプタと前記ディスクドライブの第 1 のドライブポート群が第 1 のスイッチを介して接続され、第 1 のディスクアダプタと前記ディスクドライブの第 2 のドライブポート群が第 2 のスイッチを介して接続され、第 2 のディスクアダプタと前記ディスクドライブの第 2 のドライブポート群が第 2 のスイッチを介して接続され、第 2 のディスクアダプタと前記ディスクドライブの第 1 のドライブポート群が第 1 のスイッチを介して接続され、さらに第 1 のスイッチと第 2 のスイッチが接続されたディスク装置の制御方法において、

前記第 1 のディスクアダプタあるいは第 2 のディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定することを特徴とするディスク装置の制御方法。

【書類名】 明細書**【発明の名称】 ディスク装置及びその制御方法****【技術分野】****【0001】**

本発明は、コンピュータシステムにおける2次記憶装置に関し、特に入出力データ転送性能が高いディスク装置に関する。

【背景技術】**【0002】**

現在のコンピュータシステムにおいては、CPU（中央処理装置）が必要とするデータは2次記憶装置に保存され、CPUなどが必要とするときに応じて2次記憶装置に対してデータの書き込みおよび読み出しを行う。この2次記憶装置としては、一般に不揮発な記憶媒体が使用され、代表的なものとして磁気ディスク装置や、光ディスク装置などのディスク装置がある。近年高度情報化に伴い、コンピュータシステムにおいて、この種の2次記憶装置の高性能化が要求されている。

【0003】

高性能ディスク装置では、入出力インターフェイスとしてファイバチャネルが使用されることが多い。図2.0、図2.1、図2.2にファイバチャネルの接続トポロジを示す。図2.0は、ポイントツーポイント(point to point)と呼ばれるトポロジである。このトポロジでは、ファイバチャネルのポートはN_Portと呼ばれ、一对のN_Port間は物理的には送受信2本の伝送路で相互に接続されている。図2.1は、アービトレイテッドループ(Arbitrated Loop、以下FC-ALと記す)と呼ばれるトポロジである。FC-ALにおけるファイバチャネルのポートはNL_Port(Node Loop Port)と呼ばれ、各NL_Portをループ状に接続するトポロジである。FC-ALは多数のディスクドライブを接続する場合に多く適用される。図2.2は、ファブリック(Fabric)と呼ばれるトポロジである。このトポロジは、サーバやストレージ装置のポート(N_Port)とファイバチャネルのスイッチのポート(F_Port)を接続するトポロジである。ポイントツーポイント・トポロジとファブリック・トポロジでは、接続された一对のポート間で全二重データ転送が可能である。

【0004】

図2.3と図2.4に、ファイバチャネルFCP(Fibre Channel Protocol for SCSI、以下FCPと記す)におけるエクスチェンジの例を示す。一般的に、エクスチェンジはシーケンス群からなり、またシーケンスは一連の動作を行う(又は複数の)フレームからなる。図2.3はReadの例である。イニシエータからターゲットへReadコマンドが送られ(FCP_CMND)、それに対してターゲットからイニシエータへ読み出しデータが送られ(FCP_DATA)、最後にターゲットからイニシエータへステータス情報が送信されて(FCP_RSP)、エクスチェンジが終了する。図2.4はWriteの例である。イニシエータからターゲットへWriteコマンドが送られ(FCP_CMND)、必要に応じてターゲットからイニシエータへバッファ制御情報が送られ(FCP_XFER_RDY)、それに対してイニシエータからターゲットへ書き込みデータが送られ(FCP_DATA)、最後にターゲットからイニシエータへステータス情報が送信されて(FCP_RSP)、エクスチェンジが終了する。このようにFCPでは、データ転送に方向性があり、多くの場合半二重オペレーションが行われている。データを送信しながら別のデータを平行して受信する場合を全二重オペレーションという。

【0005】

ファイバチャネルは全二重データ転送が可能なので、FCPにおいても全二重オペレーションを活用すればデータ転送能力が向上する。FCPにおいて全二重データ転送を実現する第1の従来技術としては、例えばキューロジック(QLogic)社が発行しているホワイトペーパー“Full-Duplex and Fibre Channel, http://www.qlogic.com/documents/datasheets/knowledge_data/whitepapers/tb_duplex.pdf”に記載の方法がある。第1の従来技術では、ディスクドライブを接続した複数のFC-ALとサーバを、スイッチを介して接続し、サーバと複数のFC-ALとの間で平行してデータ転送を行う。

【0006】

上位処理装置と記憶制御装置間において全二重データ転送を実現する方法が、特開 2003-851117 号の「記憶処理装置およびその運用方法」に開示されている。以下、該公報に記載の従来技術を第 2 の従来技術と呼ぶ。第 2 の従来技術では、上位装置と記憶制御装置間が全二重動作となるように、上位装置からのコマンド、データ量等に従って、入出力処理を行うチャンネルプロセッサを制御する。

【0007】

ディスクアレイ制御部とディスクドライブとをスイッチ接続したディスクアレイシステムが、特開 2000-222339 号の「ディスクサブシステム」に開示されている。以下、該公報に記載の従来技術を第 3 の従来技術と呼ぶ。

【0008】

【特許文献 1】特開 2003-851117 号の「記憶処理装置およびその運用方法」

【0009】

【特許文献 2】特開 2000-222339 号の「ディスクサブシステム」

【非特許文献 1】キューロジック(QLogic)社が発行しているホワイトペーパー“Full-Duplex and Fibre Channel, http://www.qlogic.com/documents/datasheets/knowledge_data/whitepapers/tb_duplex.pdf”

【発明の開示】

【発明が解決しようとする課題】

【0010】

ネットワーク技術の進歩に伴い、1 チャンネル当りのデータ転送速度は年々増加している。例えばディスク装置に使われるファイバチャンネルでは、現状でチャンネル当りのデータ転送速度が 1 Gbps から 2 Gbps であるが、近い将来 4 Gbps から 10 Gbps へ高速化されることが予定されている。サーバとディスク装置間（以下フロントエンドと呼ぶ）のスループットはこの高速化に従うことが予想される。ところが、ディスク装置内のディスクアダプタとディスクアレイ間（以下バックエンドと呼ぶ）のスループットは以下の理由により、フロントエンドほど高速化されないと予想される。

【0011】

第 1 の理由は、ディスクドライブは機械部品を含むので、電子、光素子のみ高速化を行えば良いフロントエンドに比べ高速化が難しいこと、である。第 2 の理由は、たとえディスクドライブが高速化したとしても、全てのディスクドライブ毎に高速インターフェイスを搭載するのは、多数のディスクドライブを有するディスク装置の高コスト化を招く、ことである。そこで、チャンネル当たりのスピードを高速化せずに、ファイバチャンネルの全二重データ転送能力を活用することにより、ディスク装置のバックエンドを高スループット化することが考えられる。

【0012】

ファイバチャンネルインターフェイスを有するディスクドライブでは、信頼性を高めるため、複数の入出力ポートを備えるディスクドライブが一般的である。第 1 の従来技術では、複数入出力ポートを有するディスクドライブについて考慮されておらず、バックエンドが複数入出力ポートを有するディスクドライブからなるディスク装置への適用は困難である。

【0013】

第 2 の従来技術では、データ転送中のダイナミックな制御が必要であり、制御方法が複雑という問題があった。また、第 2 の従来技術では、ディスク装置のバックエンドにおける全二重データ転送については触れられていない。

【0014】

第 3 の従来技術では、複数の入出力ポートを備えるディスクドライブのバックエンドへの適用と、全二重データ転送については触れられていない。

【0015】

本発明の目的は、ディスク装置のバックエンドに好適な、全二重データ転送ネットワークを有するディスク装置を提供することにある。

【0016】

本発明の他の目的は、信頼性の高いバックエンド・ネットワークを有するディスク装置を提供することにある。

【課題を解決するための手段】**【0017】**

上記目的を達成するため、本発明は、チャネルアダプタとキャッシュメモリとディスクアダプタとを有するディスクコントローラと、複数の入出力ポートを備えたディスクドライブを有するディスクアレイとからなるディスク装置において、前記ディスクアダプタと前記ディスクアレイを、スイッチを介して接続し、前記ディスクアダプタと前記ディスクドライブ間を転送されるエクスチェンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定するようにした。

【0018】

また、コマンドの種類がデータの読み出しコマンドか、あるいはデータの書き込みコマンドかによって、前記フレームの送信先ドライブポートを決定するようにした。

【0019】

さらに、データの読み出しに関する上記エクスチェンジと、データの書き込みに関する上記エクスチェンジが、平行して進行するようにした。

【0020】

さらにまた、前記ディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類によって、前記スイッチと前記ディスクドライブ間において転送されるフレームが通るパスを決定するようにした。

【0021】

また、コマンドの種類がデータの読み出しコマンドか、あるいはデータの書き込みコマンドかによって、前記スイッチと前記ディスクドライブ間において前記フレームが通るパスを決定するようにした。

【0022】

さらにまた、前記ディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類によって、前記ディスクアダプタから前記ディスクドライブへ送るフレーム内の送信先情報を前記ディスクアダプタにおいて決定し、前記スイッチは、前記ディスクアダプタが接続されたポートと前記ディスクアレイを構成するディスクドライブが接続された各ポートとのポート間接続の切り換えを、入力されたフレーム毎に、該フレーム内の送信先情報に従って行うようにした。

【0023】

また、前記スイッチは、前記ディスクアダプタが接続されたポートと前記ディスクアレイを構成するディスクドライブが接続された各ポートとのポート間接続を、入力されたフレーム毎に、前記ディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類とフレーム内の送信先情報に従って、切り換えるようにした。

【0024】

さらに、前記スイッチは、前記ディスクアダプタから前記ディスクドライブへ転送されるフレームについては該フレーム内の送信先情報と誤り検出符号を変更し、前記ディスクドライブから前記ディスクアダプタへ転送されるフレームについては該フレーム内の送信元情報と誤り検出符号を変更するようにした。

【0025】

さらにまた、前記ディスクアダプタと前記ディスクドライブの第1のドライブポート群を、第1のスイッチを介して接続し、前記ディスクアダプタと前記ディスクドライブの第2のドライブポート群を、第2のスイッチを介して接続し、さらに第1のスイッチと第2のスイッチを接続し、前記ディスクアダプタと前記ディスクドライブ間のエクスチェンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定するようにした。

【0026】

また、第 1 のディスクアダプタと前記ディスクドライブの第 1 のドライブポート群を第 1 のスイッチを介して接続し、第 1 のディスクアダプタと前記ディスクドライブの第 2 のドライブポート群を第 2 のスイッチを介して接続し、第 2 のディスクアダプタと前記ディスクドライブの第 2 のドライブポート群を第 2 のスイッチを介して接続し、第 2 のディスクアダプタと前記ディスクドライブの第 1 のドライブポート群を第 1 のスイッチを介して接続し、さらに第 1 のスイッチと第 2 のスイッチを接続し、前記第 1 のディスクアダプタあるいは第 2 のディスクアダプタと前記ディスクドライブ間のエキステンジに含まれるコマンドの種類によって、転送されるフレームの送信先ドライブの入出力ポートを決定するようにした。

【発明の効果】

【0027】

本発明によれば、容易な制御で全二重データ転送が可能なバックエンド・ネットワークを有するディスク装置が実現でき、ディスク装置のスループットを向上させる効果を奏する。

【発明を実施するための最良の形態】

【0028】

以下、図面を参照して本発明に係る実施の形態を詳細に説明する。

【0029】

(第 1 の実施の形態)

図 1 に、本発明の第 1 の実施の形態であるディスク装置の構成を示す。ディスク装置は、ディスクコントローラ D K C と、ディスクアレイ D A 1 と、スイッチ S W からなる。ディスクコントローラ D K C は、チャンネルアダプタ C H A と、キャッシュメモリ C M と、ディスクアダプタ D K A からなる。チャンネルアダプタ C H A とキャッシュメモリ C M とディスクアダプタ D K A は、相互接続網 N W で接続されている。チャンネルアダプタ C H A は、チャンネル C 1、C 2 を介して上位装置（図示せず）と接続する。ディスクアダプタ D K A はチャンネル D 0 1、D 0 2 とスイッチ S W を介して、ディスクアレイ D A 1 と接続されている。

【0030】

図 2 は、チャンネルアダプタ C H A の構成を示す。

チャンネルアダプタ C H A は、チャンネル C 1、C 2 に接続されたホストチャンネル・インターフェイス 2 1 と、相互結合網 N W に接続されたキャッシュメモリ・インターフェイス 2 2 と、サービスプロセッサ S V P に接続するためのネットワーク・インターフェイス 2 3 と、上位装置との間でのデータ転送を制御するためのプロセッサ 2 4 と、該プロセッサが参照する各種のテーブルや実行すべきソフトウェアを格納したローカルメモリ 2 5 と、これらの要素間の相互接続するプロセッサ周辺制御部 2 6 からなる。

【0031】

サービスプロセッサ S V P は、プロセッサ 2 4、3 4（後出）が参照する上記各種テーブルの設定、変更、あるいはディスク装置の稼働状況のモニタ等に使用する。

【0032】

ホストチャンネル・インターフェイス 2 1 は、チャンネルバス C 1、C 2 上のデータ転送プロトコルと、ディスクコントローラ内部のデータ転送プロトコルとの間の変換機能を有し、ホストチャンネル・インターフェイス 2 1 とキャッシュメモリ・インターフェイス 2 2 との間は信号線 2 7 によって接続されている。

【0033】

図 3 は、ディスクアダプタ D K A の構成を示す。

ディスクアダプタ D K A は、相互結合網 N W に接続されたキャッシュメモリ・インターフェイス 3 1 と、ディスクチャンネル D 0 1、D 0 2 に接続されたディスクチャンネル・インターフェイス 3 2 と、サービスプロセッサ S V P に接続するためのネットワーク・インターフェイス 3 3 と、プロセッサ 3 4 と、該プロセッサが参照する各種のテーブルや実行すべきソフトウェアを格納したローカルメモリ 3 5 と、これらの要素間の相互接続するプロ

セッサ周辺制御部 36 とからなる。

【0034】

キャッシュメモリ・インターフェイス 31 とディスクチャネル・インターフェイス 32 との間は信号線 37 によって接続されている。ディスクチャネル・インターフェイス 32 は、ディスクコントローラ内部のデータ転送プロトコルと、ディスクチャネル D01、D02 上のデータ転送プロトコル、例えば、FCP との間の変換機能を備えている。

【0035】

本実施の形態におけるディスクアレイ DA1 の構成を述べる。図 1 に示したディスクアレイ DA1 は、チャンネル D11、D12 上に接続した 4 個のディスクドライブからなるディスクアレイと、D13、D14 上に接続した 4 個のディスクドライブからなるディスクアレイからなる。チャンネル D11 を例にとると、ディスクドライブ DK0、DK1、DK2、DK3 が、チャンネル D11 上に接続されている。このように、多数のドライブを一つのチャンネル上に接続してディスクドライブにアクセスする方法としては、ファイバチャネル・アービトレイテッド・ループ（以下 FC-AL と呼ぶ）がある。

【0036】

図 4 に本実施の形態における FC-AL の接続形態の詳細を示す。各ディスクドライブは、それぞれ NL ポートを 2 個有する。各ディスクドライブの入出力ポートおよびスイッチ SW の入出力ポートは、送信機 Tx と受信機 Rx を有する。スイッチ SW のディスクアレイ DA1 接続側入出力ポートは、FL (Fabric Loop) ポートである。チャンネル D11 により、スイッチ SW、ディスクドライブ DK0、DK1、DK2、DK3 をループ状に接続する。同様にチャンネル D12 により、スイッチ SW、ディスクドライブ DK0、DK1、DK2、DK3 をループ状に接続する。これら 2 個のループは、ファイバチャネルのパブリックループであり、ディスクドライブ DK0、DK1、DK2、DK3 は、スイッチ SW を介して、ディスクアダプタ DKA のディスクチャネル・インターフェイス 32 と通信可能である。以上、チャンネル D11、D12 の接続形態を例に説明したが、チャンネル D13、D14 でも同様である。

【0037】

次に、本実施の形態のスイッチ動作を説明する。図 5 に示すように、スイッチ SW は入出力ポート P1、P2、P3、P4、P5、P6 を有する。ポート P1、P2、P3、P4、P5、P6 は、全二重データ転送可能な入出力ポートである。動作の例として、ポート P1 からフレームを入力し、ポート P2、P3、P4、P5、P6 から出力する場合を述べる。スイッチ SW は図 5 に示すように、クロスバスイッチ 510 と、スイッチコントローラ 511 からなる。クロスバスイッチ 510 は、この例では 6×6 のクロスバスイッチであり、入力ポート in1、in2、in3、in4、in5、in6 と、出力ポート out1、out2、out3、out4、out5、out6 を有する。

【0038】

ポート P1 から入力したフレームは、シリアルパラレル変換装置 SP1 と、バッファメモリ BM1 と、8b10b 変換デコード DEC1 とフレームヘッダ解析部 501 を経由し、スイッチコントローラ 511 と入力ポート in1 へ入力される。スイッチコントローラ 511 において、入力フレームのヘッダ部分に書かれた送信先ポート ID に従い、クロスバスイッチ 510 を切り換える。例として、ポート P6 に接続されたデバイスのポートが送信先として選ばれた場合は、入力したフレームは出力ポート out6 と、8b10b 変換エンコーダ ENC1 と、バッファメモリ BM2 と、パラレルシリアル変換装置 PS1 を経由して、ポート P6 から出力される。ここで、バッファメモリ BM1、BM2 は FIFO (First-In First-out) メモリである。

【0039】

ディスクアダプタ DKA とディスクアレイ DA1 を、スイッチ SW を介して接続したことにより、ディスクアダプタ DKA はディスクドライブ DK0～DK7 の任意の入出力ポートにフレームを送信することができる。

【0040】

図 1 において、ディスクアダプタ D K A とスイッチ S W は、2 本のチャンネル D 0 1、D 0 2 によって接続されているが、説明を容易にする為、ここではチャンネル D 0 1 のみ使用する場合を考える。ディスクアダプタ D K A 内のプロセッサ 3 4 が参照する、バックエンド管理テーブルの例を図 6 に示す。ドライブ番号に対応して、Read コマンドの送信先ドライブのポート I D と、Write コマンドの送信先ドライブポート I D が、図 6 の 6 0 1 にセットされている。6 0 1 において、PID_0.a~PID_7.a は、チャンネル D 1 1 あるいはチャンネル D 1 3 と接続された F C - A L 上のディスクドライブのポート I D を示す。PID_0.b~PID_7.b は、チャンネル D 1 2 あるいはチャンネル D 1 4 と接続された F C - A L 上のディスクドライブのポート I D を示す。通常動作時(各ドライブポートが正常)は、ディスクアダプタ D K A から送信された Read コマンドは、チャンネル D 0 1 とスイッチ S W を経由し、送信先ポートである PID_0.a~PID_7.a のいずれかに届く。読み出しデータは、Read コマンドと方向は逆であるが同じパスを通る。一方 Write コマンドと書き込みデータは、チャンネル D 0 1 とスイッチ S W を経由して、送信先ポートである PID_0.b~PID_7.b のいずれかに届く。

【 0 0 4 1 】

例として、ドライブ番号 0 のディスクドライブに対する Read と、ドライブ番号 1 のディスクドライブに対する Write の動作について述べる。図 3 のプロセッサ 3 4 は、図 6 の 6 0 1 を参照することにより、Read コマンドを PID_0.a へ、Write コマンドを PID_1.b へ送信する。Read コマンドが通るパスは、ディスクアダプタ D K A から始まり、チャンネル D 0 1、スイッチ S W、チャンネル D 1 1 を経由し、ポート PID_0.a へ至る。Write コマンドが通るパスは、ディスクアダプタ D K A から始まり、チャンネル D 0 1、スイッチ S W、チャンネル D 1 2 を経由し、ポート PID_1.b へ至る。このようにコマンドの種類(Read/Write)によって、スイッチとディスクアレイ間でデータが通るパスを分けているので、Read のエクステンジと Write のエクステンジを平行して進行させることができる。

【 0 0 4 2 】

図 2 5 は、Read エクステンジと Write エクステンジが平行して進行している場合の、ディスクアダプタ D K A - スwitch 間 S W (チャンネル D 0 1 上)のフレーム例を示す図である。Read エクステンジと Write エクステンジのデータ転送シーケンスに重なる時間の部分が生じるように、ディスクアダプタ D K A は Read コマンドと Write コマンドを発行する。ディスクアダプタ D K A は、Read コマンドと Write コマンドを、必ずしも同時に発行する必要はない。また、Read エクステンジと Write エクステンジで、データ転送サイズが同じである必要もない。さらに、複数の Read エクステンジと複数の Write エクステンジが平行して進行してもよい。

【 0 0 4 3 】

このとき、チャンネル D 0 1 上では平行して双方向にデータ転送が行われる。つまり、ディスクアダプタ D K A とスイッチ S W 間において全二重オペレーションの状態となる。プロセッサ 3 4 は、Read エクステンジと Write エクステンジのデータ転送シーケンスの時間が重なるようにコマンドを発行することにより、ディスクアダプタ D K A とスイッチ S W 間の全二重オペレーションが可能となる。ディスクアダプタ D K A が送信先ポート I D を決定するためには、エクステンジ開始時に 1 回だけ管理テーブルを参照するだけでよいので、非常に容易な手段で全二重オペレーションを実現できることが分かる。

【 0 0 4 4 】

もし、ディスクドライブのポートの一方に障害が発生した場合は、6 0 2 あるいは 6 0 3 の設定を適用することにより、ディスクアダプタ D K A はディスクアレイ D A 1 にアクセス可能である。例えば、ドライブ番号 2 のディスクドライブを Read しようとしたが、PID_2.a のポートに障害が発生した場合を考える。このとき、プロセッサ 3 4 は 6 0 2 の設定を参照し、ドライブ番号 2 のディスクドライブに対しては Read コマンドを PID_2.b のポートへ送信する。同様に、ドライブ番号 3 のディスクドライブを Write しようとしたが、PID_3.b のポートに障害が発生した場合を考える。このとき、プロセッサ 3 4 は 6 0 3 の設定を参照し、ドライブ番号 3 のディスクドライブに対しては Write コマンドを PID_3.a のポ

ートへ送信する。

【0045】

バックエンド管理テーブルの他の例を図7に示す。図6の管理テーブルとの違いは、例えば701の項目に記されているように、Readコマンド送信先ポートとWriteコマンド送信先ポートが同じFC-A L上に設定されている点である。この場合、Readエクステンジとriteエクステンジが同じFC-A Lの帯域を共有することになる。しかし、例えば、ドライブ番号0のディスクドライブへのReadとドライブ番号2のディスクドライブへのWriteのように、異なるFC-A L上のエクステンジを平行して進行させることにより、チャンネルD01上では平行して双方向にデータ転送が行われる。このように同じFC-A L上にReadのエクステンジとWriteのエクステンジが混在する設定にしている、全二重オペレーションは問題なく実施でき、半二重オペレーション実行時に対してスループットが向上する。

【0046】

以上、第1の実施の形態においては、ディスクアダプタが、発行するコマンドの種類(Read/Write)に応じて、送信先のディスクポートを決定していた。同様の処理はスイッチにおいても実施可能である。

【0047】

(第2の実施の形態)

図8から図11A、図11Bに、本発明における第2の実施の形態を示す。本実施の形態では、ディスクアダプタで設定された送信先ドライブポートにかかわらず、スイッチにおいてフレーム内情報を変更することで全二重オペレーションを実現する。

【0048】

図8は本実施の形態で使用するスイッチの構成を示す。図5のスイッチに対して、メモリ812が追加され、スイッチ部810は共有メモリ型となっている。プロセッサ811は、共有メモリ型スイッチ810に蓄えられたフレームを読み書き可能である。メモリ812には図11Aおよび図11Bに示す管理テーブルが格納され、プロセッサ811は図10のフローチャートに従い、フレーム変更処理を行う。図11Aの管理テーブルでは、ディスクアダプタからスイッチへ送信されたフレーム内の送信先ポートID 1101に対する変更ポートID 1102、1103が対応づけられている。1102はReadのエクステンジに対する変更ポートIDであり、1103はWriteのエクステンジに対する変更ポートIDである。図11Bの管理テーブルは、エクステンジ毎に作製される変更項目を示し、図10のフローチャートに従って作製、参照される。

【0049】

図10のフローチャートは、スイッチをフレームが通過するたびに実行される。詳細には、このフレーム変更処理は、ディスクアダプタとスイッチ間での入出力に際して実行される。二重実行を防ぐため、スイッチとディスクアレイとの入出力に際しては、実行されない。

【0050】

ステップ1001において、入力フレームがFCP_CMNDかどうかチェックし、新しいエクステンジが開始したかを判断する。次にフレームがFCP_CMNDの場合は、ステップ1002においてコマンド種類を調べる。次にコマンドがReadまたはWriteの場合は、ステップ1003に進む。

【0051】

ステップ1003において、FCP_CMNDフレームからエクステンジID OX_IDと、送信先ID D_ID、送信元ID S_IDを読み取る。読み取ったOX_ID、S_ID、D_IDを、それぞれ図11Bテーブルの1104、1105、1106にセットする。1106にセットした送信先ポートIDと図11Aテーブルから、変更後の送信元ポートID 1107、送信先ポートID 1108をセットする。一つのエクステンジにおいて、ディスクアダプタからスイッチへ入力されるフレームに対しては1109の変更を行い、スイッチからディスクアダプタへ出力されるフレームに対しては1110の変更を行い、合わせて2種類のフレーム変更処理を行う。1109では送信

先ポート I D のみ変更し、1110では送信元ポート I D のみ変更する。1110の変更は、ディスクアダプタへ送信されるフレームの S_ID と D_ID に矛盾を発生させないために必要である。

【0052】

次に図 10 のステップ 1004 に進む。ここで先に作製したテーブル図 11 B に従い、フレームの送信先ポート I D D_ID を変更し、さらに C R C (Cyclic Redundancy Check) を計算し直してフレーム内の所定の位置に書き戻す。

【0053】

ステップ 1001 で判定が No の場合は、ステップ 1005 へ進む。フレーム内のエクスチェンジ I D OX_ID、送信元ポート I D S_ID、送信先ポート I D D_ID を読み取り、図 11 B テーブルの各行と比較する。テーブル内に該当する項目があれば（各行において OX_ID、S_ID、D_ID が全て一致する項目があれば）、ステップ 1006 へ進み、図 11 B テーブルに従って送信元ポート I D S_ID、送信先ポート I D D_ID を変更し、さらに C R C を計算し直してフレーム内の所定の位置へ書き戻す。次にステップ 1007 へ進み、エクスチェンジ終了条件を調べる。もしエクスチェンジが終了した場合には、ステップ 1008 へ進み、図 11 B から該当するエクスチェンジの行のエントリを削除する。

【0054】

なお、図 9 にフレームの構造（例として FCP_CMND）を示すが、送信先ポート I D は 901、送信元ポート I D は 902、エクスチェンジ I D は 903、コマンド種類は 904、エラー検出情報は 905、エクスチェンジ終了状況は 906 を調べることにより、容易に知ることができる。

【0055】

以上述べたように本実施の形態によれば、スイッチにおけるフレーム変更処理により、第 1 の実施の形態と同様の動作を実現できるので、ディスクアダプタの負荷を低減できる効果がある。

（第 3 の実施の形態）

図 12 に本発明の、第 3 の実施の形態であるディスク装置の構成を示す。本実施の形態のディスク装置は、スイッチを二重化した点に特徴がある。本実施の形態において、ディスクアダプタ DKA と、スイッチ SW1、SW2 と、ディスクアレイ DA2 との間のデータ転送方式は、ファイバチャネルを使用している。

【0056】

本実施の形態のディスク装置は、ディスクコントローラ DKC と、スイッチ SW1、SW2 と、ディスクアレイ DA2 からなる。ディスクコントローラ DKC は、チャネルアダプタ CHA と、キャッシュメモリ CM と、ディスクアダプタ DKA からなる。

【0057】

ディスクアダプタ DKA とスイッチ SW1 をチャネル D01 で接続し、ディスクアダプタ DKA とスイッチ SW2 をチャネル D02 で接続する。スイッチ SW1 とスイッチ SW2 は、チャネル 1201 で接続する。

【0058】

ディスクアレイ DA2 を構成する各ディスクドライブは、入出力ポートを 2 個有する。例えば、ディスクドライブ DK0、DK4、DK8、DK12 は、チャネル D11 および D21 の両チャネルと接続する。ディスクアレイ DA2 は、チャネル D11 と D21 に接続した 4 個のディスクからなるディスクアレイと、D12 と D22 に接続した 4 個のディスクからなるディスクアレイと、D13 と D23 に接続した 4 個のディスクからなるディスクアレイと、D14 と D24 に接続した 4 個のディスクからなるディスクアレイ、からなる。チャネル D11、D12、D13、D14、D21、D22、D23、D24 は、FC-AL でディスクドライブを接続する。

【0059】

本実施の形態におけるバックエンド管理テーブル例を図 13 に示す。項目 1301 (VDEV) は、ディスクドライブが属する論理グループを示す。ディスクアダプタ DKA は、項

目1302、1303、1304のDKA Port値が0の場合はチャンネルD01を、値が1の場合はチャンネルD02を使用して、スイッチSW1あるいはスイッチSW2と接続し、ディスクアレイDA2と通信する。PID_0.a~PID_15.aは、スイッチSW1に接続されたFC-A L上のディスクドライブのポートIDを示す。PID_0.b~PID_15.bは、スイッチSW2に接続されたFC-A L上のディスクドライブのポートIDを示す。通常動作時（SW1、SW2共に正常）は、ディスクアダプタDKAから送信されたReadコマンドはSW1を経由し、送信先ポートであるPID_0.a~PID_15.aのいずれかに届く。読み出しデータは、Readコマンドと方向は逆であるが同じパスを通る。一方Writeコマンドと書き込みデータは、スイッチSW1と、チャンネル1201と、スイッチSW2を経由して、送信先ポートであるPID_0.b~PID_15.bのいずれかに届く。

【0060】

例として、ドライブ番号0のディスクドライブに対するReadと、ドライブ番号4のディスクドライブに対するWriteの動作について述べる。Readコマンドが通るパスは、ディスクアダプタDKAから始まり、チャンネルD01、スイッチSW1、チャンネルD11を経由し、ポートPID_0.aへ至る。Writeコマンドが通るパスは、ディスクアダプタDKAから始まり、チャンネルD01、スイッチSW1、チャンネル1201、スイッチSW2、チャンネルD21を経由し、ポートPID_4.bへ至る。

【0061】

このようにコマンドの種類(Read/Write)によって、スイッチとディスクアレイ間でデータが通るパスを分けているので、ReadのエクステンジとWriteのエクステンジを平行して進行させることができ、ディスクアダプタDKAとスイッチSW1間の全二重オペレーションが可能となる。

【0062】

スイッチSW1に障害が発生した場合は、1303の設定を適用し、SW2に障害が発生した場合は1304の設定を適用することにより、1個のスイッチに障害が発生した場合でもディスクアダプタDKAは、ディスクアレイDA2にアクセス可能である。ただし、スイッチ障害発生時は、一つのFC-A Lの帯域を共有するコマンド数が増えることになるので、正常時に比べスループットは低下する場合がある。

【0063】

図14A、図14B、図14C及び図15により、本発明の実施の形態におけるスループット向上効果を説明する。図14A、図14B、及び図14Cに比較検討したトポロジを示す。図14A、図14B、及び図14Cは、FC-A Lに接続された4台のディスクドライブに対して、内ドライブ2個に対してはWrite、他のドライブ2個に対してはReadを実行するトポロジを示している。図14Aは従来のディスク装置のトポロジである。ディスクアダプタにFC-A Lが直結されている。ループ速度は1Gbpsである。図14Bは、本実施の形態のトポロジであるが、コマンド種類(Read/Write)毎にループを分けた例である。ループ速度は1Gbpsであるが、ディスクアダプタとスイッチ間および2個のスイッチ間のチャンネル速度は2Gbpsとした。図14Cは、図14B同様、本実施の形態のトポロジであるが、異なるコマンド(Read/Write)を同じループに混在させた例である。ループ速度は1Gbpsであるが、ディスクアダプタとスイッチ間および2個のスイッチ間のチャンネル速度は2Gbpsとした。

【0064】

図15は、図14A、図14B、及び図14Cで示したトポロジにおけるスループットの実測例である。図15において、曲線(A)は図14A、曲線(B)は図14B、曲線(C)は図14Cのトポロジにおけるスループット特性を示す。横軸は1コマンド当たりのデータ転送サイズ(KB)を、縦軸はスループット(MB/s)である。グラフから明らかな様に、データ転送サイズ8KB以上で従来の(A)トポロジに比べてスループット向上が認められ、データ転送サイズ16KB以上で36%、データ転送サイズが128KB以上の領域では87%のスループット向上を確認できた。

【0065】

また、曲線(B)と(C)を比較するとわかるが、コマンド(Read/Write)によって使用するループを別にした方が、ループ内に異なるコマンドを混在させるよりもスループット向上効果が高い。

【0066】

以上、第3の実施の形態では、ディスクアダプタの入出力ポート2個の内1個を定常系、他1個を障害時用の交替系として説明したが、もちろん、2個の入出力ポートを同時に使用しても構わない。図16に、ディスクアダプタの2個の入出力ポートを同時使用する場合のバックエンド管理テーブル例を示す。図16の項目1601に示すように、ディスクドライブによって使用するディスクアダプタのポートを変更している。これにより、2個のディスクアダプタ・ポート間でバックエンド・ネットワーク負荷を分散できる。また、障害発生時に使用してみて初めて、交替系の故障が判明するのを防止する効果がある。

(第4の実施の形態)

図17に、本発明の第4の実施の形態であるディスク装置の構成を示す。本実施の形態において、ディスクアダプタDKA1、DKA2と、スイッチSW1、SW2と、ディスクアレイDA3との間のデータ転送方式は、ファイバチャネルを使用している。本実施の形態は、第3の実施の形態に比べてディスクコントローラを構成する要素が二重化されており、信頼性が高い特徴を有する。チャネルアダプタCHA1、CHA2、キャッシュメモリCM1、CM2、ディスクアダプタDKA1、DKA2は、2つの相互結合網NW1、NW2で相互接続されている。ディスクアダプタDKA1は、スイッチSW1あるいはSW2を介してディスクアレイと接続可能である。ディスクアダプタDKA2も同様に、スイッチSW1あるいはSW2を介してディスクアレイDA3と接続可能である。本実施の形態におけるバックエンド管理テーブルの例を、図18に示す。PID_0.a~PID31.aは、スイッチSW1に接続されたFC-A L上のディスクドライブのポートIDを示す。また、PID_0.b~PID31.bは、スイッチSW2に接続されたFC-A L上のディスクドライブのポートIDを示す。ディスクアダプタDKA1は、DKAPort値が0の場合はチャネルD01を、値が1の場合はチャネルD02を使用して、スイッチSW1あるいはSW2と接続し、ディスクアレイDA3と通信する。ディスクアダプタDKA2は、DKAPort値が0の場合はチャネルD03を、値が1の場合はチャネルD04を使用して、スイッチSW1あるいはSW2と接続し、ディスクアレイDA3と通信する。図18では、図16の管理テーブルに比べてDKA番号1801の項目が追加されている。項目1801は、二重化されているディスクアダプタのどちらを使用するかを示す。例えばDKA番号が0ならば、そのディスクドライブはディスクアダプタDKA1からアクセスする。逆に、DKA番号が1ならば、そのディスクドライブはディスクアダプタDKA2からアクセスする。ディスクアダプタのうちどちらか一方に障害が発生した場合は、他方のディスクアダプタからアクセスするように管理テーブルのDKA番号1801を変更する。本実施の形態によれば、ディスクアダプタを二重化しているので信頼性が向上できる効果の他、正常動作時は2個のディスクアダプタ間で負荷を分散できる効果がある。また、言うまでもないが、第1から第3の実施の形態と同様に、ディスクアダプタが発行するコマンドの種類によって、フレーム送信先のディスクドライブ・ポートを決定することにより、全二重オペレーション実行時のスループットが向上する効果がある。

【0067】

図18の管理テーブルでは、スイッチSW1と接続されたディスクドライブのポートはRead用、スイッチSW2と接続されたディスクドライブのポートはWrite用に割り当てている(スイッチSW1、SW2正常時)。例えば、ディスクアダプタDKA1からドライブ0へのWriteデータは、ディスクアダプタDKA1からスイッチSW1、チャネル1701、スイッチSW2を順に経由し、ドライブ0へ転送される。また、ドライブ4からディスクアダプタDKA2へのReadデータは、ドライブ4からスイッチSW1、チャネル1701、スイッチSW2を順に経由して、ディスクアダプタDKA2へ転送される。図18の設定では、スイッチ間をつなぐチャネル1701上のデータ転送方向は、スイッチSW1からスイッチSW2への一方向となる。

【0068】

第4の実施の形態における、バックエンド管理テーブルの他の例を図26に示す。図26の設定では、同じスイッチに接続するディスクドライブのポートでも、属するループによってRead用ポートあるいはWrite用ポートの割り当てを変更している点に特徴がある。

【0069】

図26によれば、ドライブ0、4、8、12・・・28およびドライブ2、6、10、14・・・30のスイッチSW1へ接続するポートはRead用ポート、スイッチSW2へ接続するポートはWrite用ポートとして割り当てる。一方、ドライブ1、5、9、13・・・29およびドライブ3、7、11、15・・・31のスイッチSW1へ接続するポートはWrite用ポート、スイッチSW2へ接続するポートはRead用ポートとして割り当てる。例えば、ドライブ0へのWriteデータは、ディスクアダプタDKA1からスイッチSW1、チャンネル1701、スイッチSW2を順に経由し、ドライブ0へ転送される。それに対して、ドライブ1からのReadデータは、ドライブ1からスイッチSW2、チャンネル1701、スイッチSW1を順に経由して、ディスクアダプタDKA1へ転送される。このように、同じスイッチに接続したドライブポートにおいて、Read/Writeコマンドとドライブポートとの対応をループ毎に変えることにより、スイッチ間で双方向にデータが流れるようになるので、チャンネル1701上においても全二重オペレーションが実現可能である。図18の設定に比べて図26の設定では、スイッチ間を接続するチャンネル1701の配線本数を節約することができる。

【0070】

(第5の実施の形態)

図19に、本発明の第5の実施の形態であるディスク装置の構成を示す。上記第1から第4の実施の形態では、バックエンド・ネットワークはファイバチャネルを使用していたが、本実施の形態ではSAS (Serial Attached SCSI)を使用した例を示す。ディスクアダプタDKA1は、エキスパンダ1904あるいはエキスパンダ1905を介してディスクアレイと接続可能である。ディスクアダプタDKA2も同様に、エキスパンダ1904あるいはエキスパンダ1905を介してディスクアレイと接続可能である。ディスクアダプタDKA1とエキスパンダ1、2間と、ディスクアダプタDKA2とエキスパンダ1、2間、および、エキスパンダ1、2間は、Wideポートで接続し、エキスパンダと各ディスクドライブ間を、Narrowポートで接続している。エキスパンダは、ファイバチャネルのスイッチに相当するが、ループ接続はサポートされていない。従って、多数のディスクドライブを接続する場合は、複数のエキスパンダを多段に接続し、ドライブ接続ポート数を増やしても良い。使用できるディスクドライブは、2ポートを有するSASドライブ1901の他、SATA (Serial ATA) ドライブ1902も接続できる。ただし、入出力ポートが1個のSATAドライブ1903の場合はセクタ1906を介して、エキスパンダ1904およびエキスパンダ1905と接続する。本実施の形態によれば、ファイバチャネルドライブより安価なSASドライブ、SATAドライブを使用できるので、ディスク装置を低コスト化できる。また、言うまでもないが、第1から第4の実施の形態と同様に、ディスクアダプタが発行するコマンドの種類によって、フレーム送信先のディスクドライブ・ポートを決定することにより、全二重オペレーション実行時のスループットが向上する効果がある。

【0071】

さらに本実施例の形態によれば、ディスクドライブの2個の入出力ポートを定常的に使用することにより全二重データ転送を実現しているので、障害発生時に初めて交替用ディスクドライブポートの障害が発見されるのを防止することができる。そして、ディスクアダプタとディスクアダプタ間の接続を2個のエキスパンダを冗長化しているため、バックエンドネットワークの信頼性が高い。

【図面の簡単な説明】**【0072】**

【図1】 本発明の第1の実施の形態のディスク装置を示す図である。

【図2】 本発明のチャンネルアダプタ10の構成例を示す図である。

- 【図 3】 本発明のディスクアダプタ 20 の構成例を示す図である。
 【図 4】 本発明のバックエンド構成例の詳細を示す図である。
 【図 5】 本発明に用いるスイッチの構成例を示す図である。
 【図 6】 本発明のディスクアダプタで参照される管理テーブル例を示す図である。
 【図 7】 本発明のディスクアダプタで参照される管理テーブル例を示す図である。
 【図 8】 本発明の第 2 の実施の形態に用いるスイッチの構成を示す図である。
 【図 9】 本発明の FCP_CMND のフレーム例を示す図である。
 【図 10】 本発明のスイッチにおける処理フロー例を示す図である。
 【図 11A】 本発明のスイッチで参照される管理テーブル例を示す図である。
 【図 11B】 本発明のスイッチで参照される管理テーブル例を示す図である。
 【図 12】 本発明の第 3 の実施の形態のディスク装置を示す図である。
 【図 13】 本発明の第 3 の実施の形態で参照される管理テーブルを示す図である。
 【図 14A】 本発明の第 3 の実施の形態の効果を説明するための比較図である。
 【図 14B】 本発明の第 3 の実施の形態の効果を説明するための図である。
 【図 14C】 本発明の第 3 の実施の形態の効果を説明するための図である。
 【図 15】 本発明の第 3 の実施の形態の効果を説明するための図である。
 【図 16】 本発明の第 3 の実施の形態で参照される管理テーブルを示す図である。
 【図 17】 本発明の第 4 の実施の形態のディスク装置を示す図である。
 【図 18】 本発明の第 4 の実施の形態で参照される管理テーブルを示す図である。
 【図 19】 本発明の第 5 の実施の形態のディスク装置を示す図である。
 【図 20】 ポイントツーポイント・トポロジを説明する図である。
 【図 21】 アービトレイテッドループ・トポロジを説明する図である。
 【図 22】 ファブリック・トポロジを説明する図である。
 【図 23】 Read動作時のエクステンションを説明する図である。
 【図 24】 Write動作時のエクステンションを説明する図である。
 【図 25】 本発明の同時進行するReadエクステンションとWriteエクステンション例を説明する図である。

【図 26】 本発明のバックエンド管理テーブル例を示す図である。

【符号の説明】

【0073】

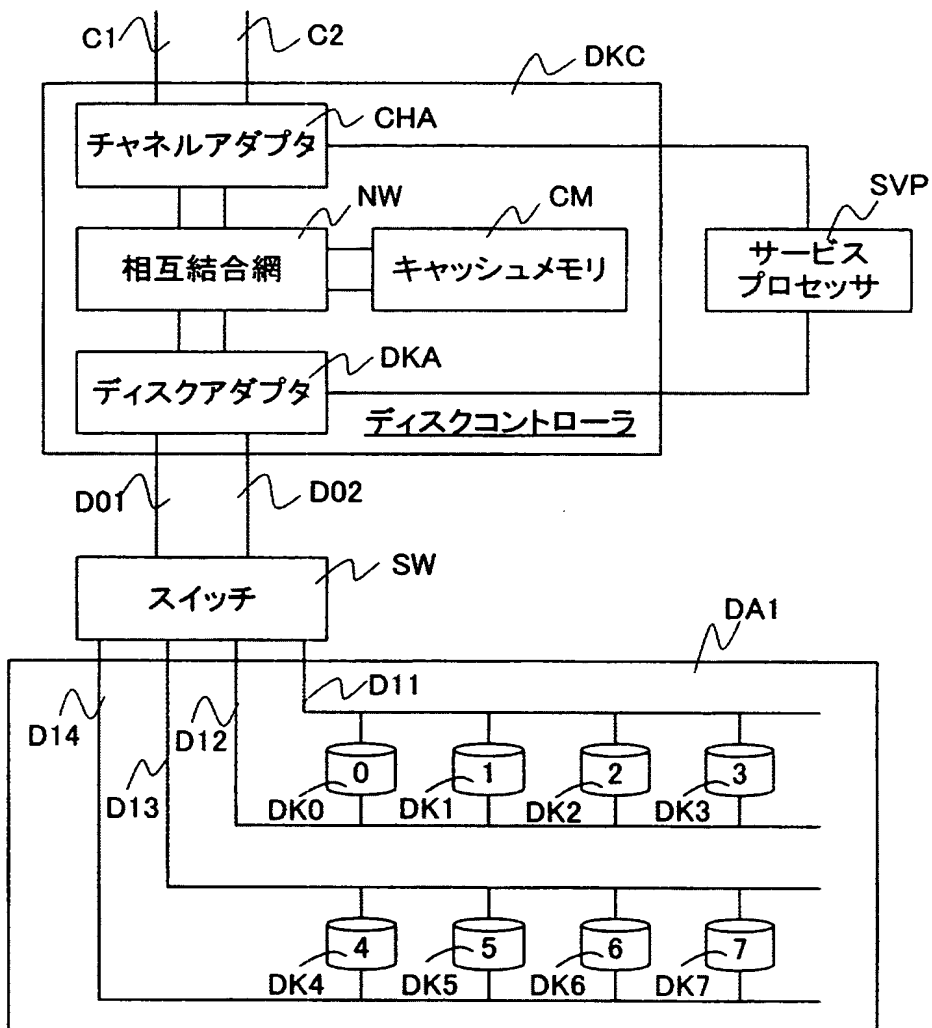
DKC・・・ディスクコントローラ、
 CHA、CHA1、CHA2・・・チャネルアダプタ、
 CM、CM1、CM2・・・キャッシュメモリ、
 DKA、DKA1、DKA2・・・ディスクアダプタ、
 DA1、DA2、DA3・・・ディスクアレイ、
 DK0、DK1、DK2、DK3、
 DK4、DK5、DK6、DK7、
 DK8、DK12・・・ディスクドライブ、
 C1、C2、D01、D02、D03、D04、D11、D12、D13、D14D2
 1、D22、D23、D24・・・チャネル
 SW、SW1、SW2・・・スイッチ
 P1、P2、P3、P4、P5、P6・・・スイッチポート
 510・・・クロスバススイッチ、
 511・・・スイッチコントローラ、
 in1、in2、in3、in4、in5、in6・・・クロスバススイッチ入力ポート
 、
 out1、out2、out3、out4、out5、out6・・・クロスバスス
 witch出力ポート、
 SP1・・・シリアルパラレル変換装置、
 PS1・・・パラレルシリアル変換装置、

BM1、BM2・・・パッファメモリ、
DEC1・・・8b10b変換デコード、
ENC1・・・8b10b変換エンコード、
Tx・・・送信機、
Rx・・・受信機、
NL・・・NLポート、
FL・・・FLポート、
N・・・Nポート、
F・・・Fポート、
NW、NW1、NW2・・・相互結合網、
VEDV0、VDEV1、VDEV2、VDEV3、VEDV4、VDEV5、V D
EV6、VDEV7・・・論理グループ、
SVP・・・サービスプロセッサ、
21・・・ホストチャネル・インターフェイス、
22、31・・・キャッシュメモリ・インターフェイス
23、33・・・ネットワーク・インターフェイス、
24、34・・・プロセッサ、
26、36・・・プロセッサ周辺制御部、
25、35・・・ローカルメモリ、
27、37・・・信号線、
810・・・共有メモリ型スイッチ、
811・・・プロセッサ、
812・・・メモリ、
1901・・・SASドライブ、
1902、1903・・・SATAドライブ、
1906・・・セクタ、
1904、1905・・・エキスパンダ。

【書類名】 図面

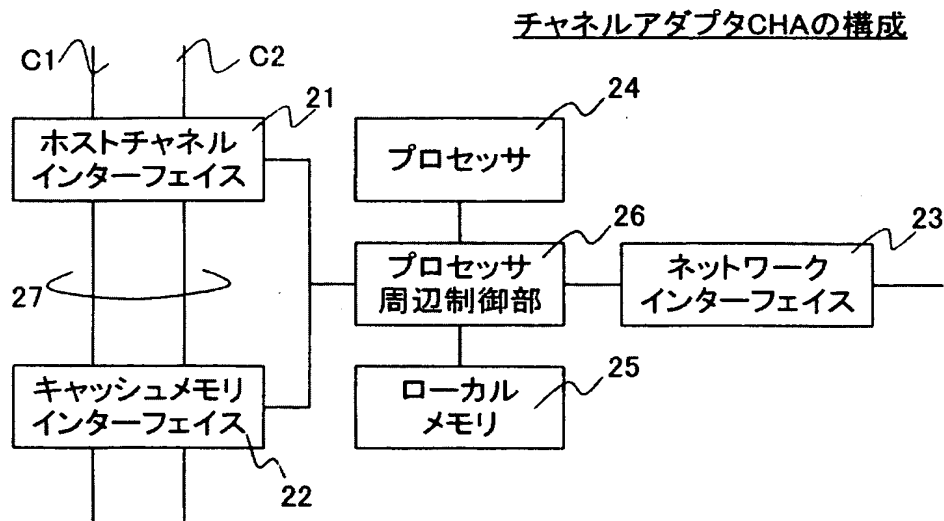
【図 1】

図1



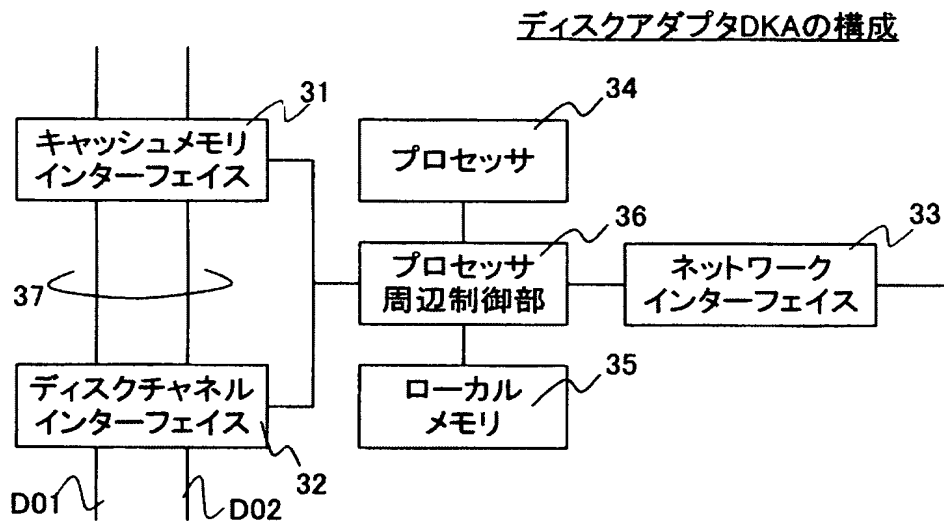
【図 2】

図2



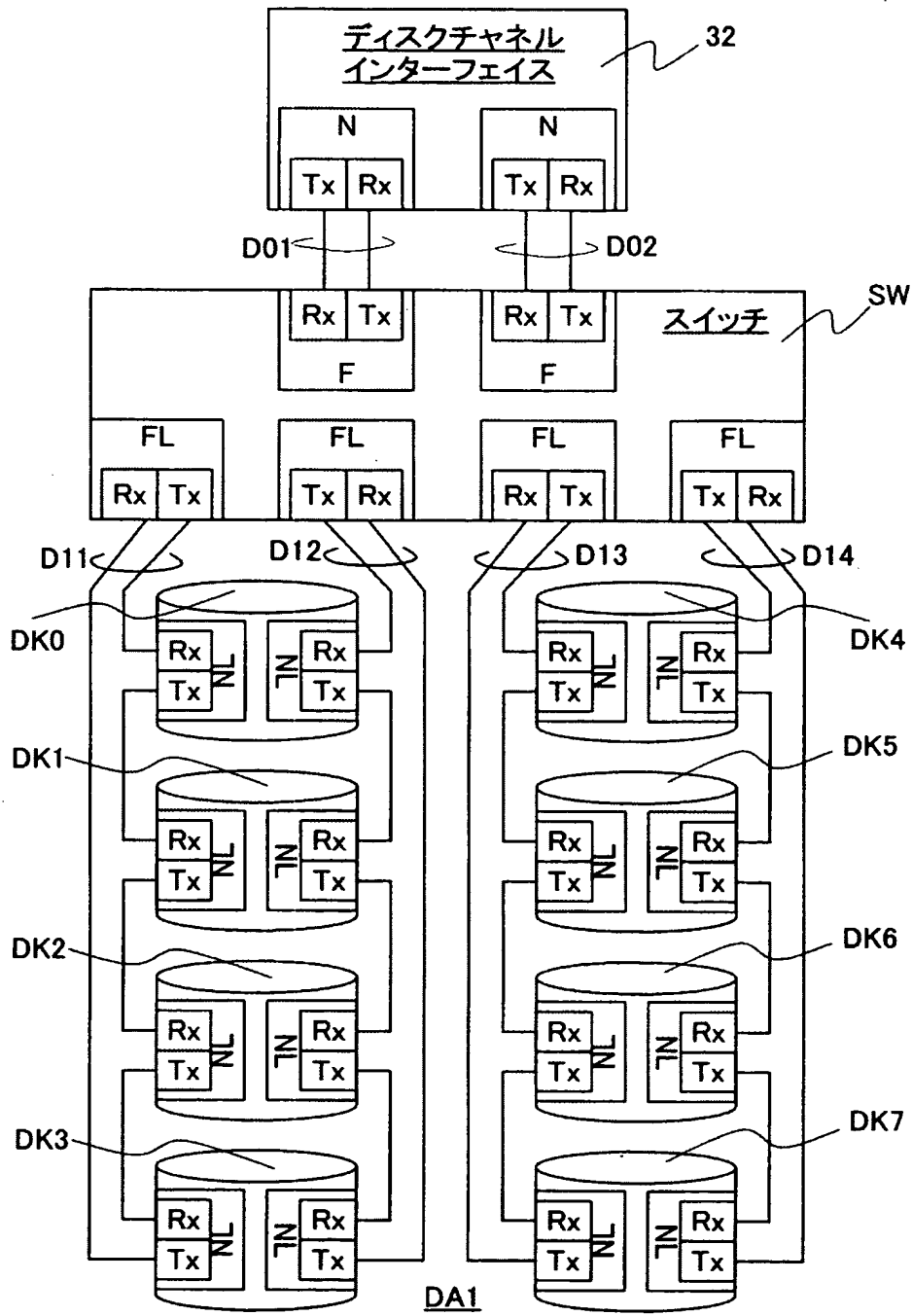
【図 3】

図3



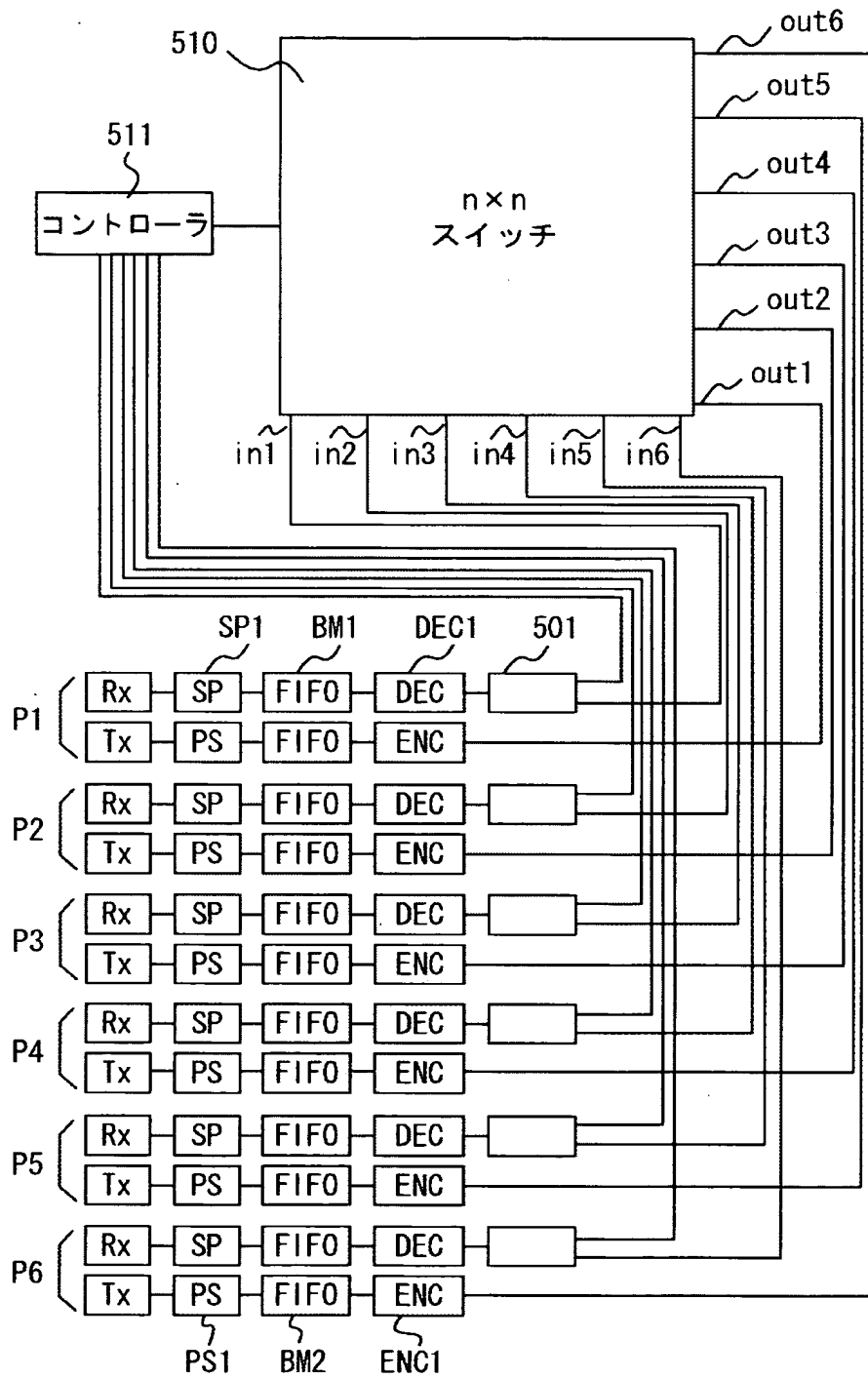
【図 4】

図4



【図 5】

図 5



【図 6】

図 6

ドライブ 番号	ドライブ正常		ポートa 障害		ポートb 障害	
	Read	Write	Read	Write	Read	Write
0	PID_0.a	PID_0.b	PID_0.b		PID_0.a	
1	PID_1.a	PID_1.b	PID_1.b		PID_1.a	
2	PID_2.a	PID_2.b	PID_2.b		PID_2.a	
3	PID_3.a	PID_3.b	PID_3.b		PID_3.a	
4	PID_4.a	PID_4.b	PID_4.b		PID_4.a	
5	PID_5.a	PID_5.b	PID_5.b		PID_5.a	
6	PID_6.a	PID_6.b	PID_6.b		PID_6.a	
7	PID_7.a	PID_7.b	PID_7.b		PID_7.a	

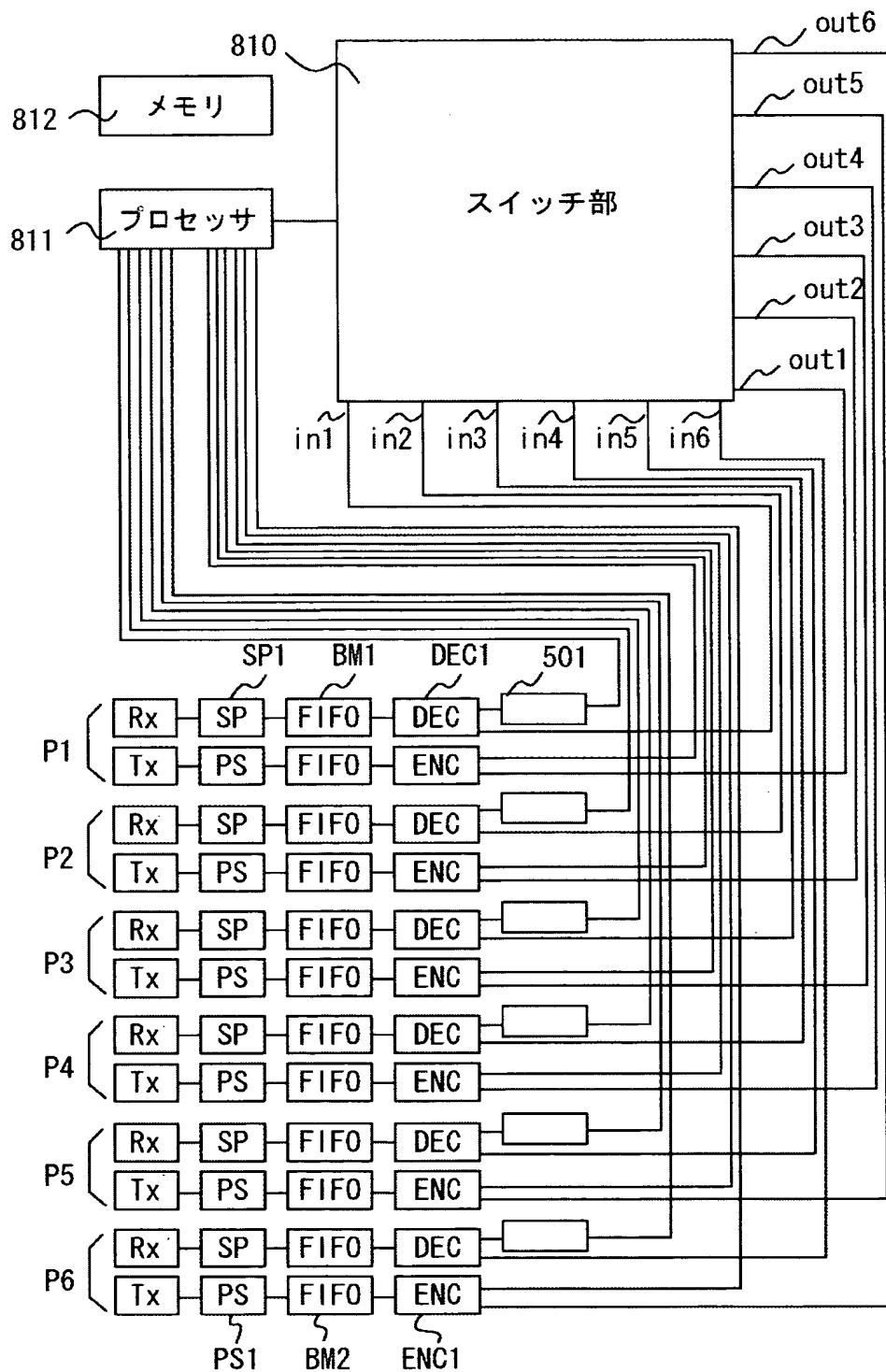
【図 7】

図 7

ドライブ 番号	ドライブ正常		ポートa 障害		ポートb 障害	
	Read	Write	Read	Write	Read	Write
0	PID_0.a	PID_0.b	PID_0.b		PID_0.a	
1	PID_1.b	PID_1.a	PID_1.b		PID_1.a	
2	PID_2.a	PID_2.b	PID_2.b		PID_2.a	
3	PID_3.b	PID_3.a	PID_3.b		PID_3.a	
4	PID_4.a	PID_4.b	PID_4.b		PID_4.a	
5	PID_5.b	PID_5.a	PID_5.b		PID_5.a	
6	PID_6.a	PID_6.b	PID_6.b		PID_6.a	
7	PID_7.b	PID_7.a	PID_7.b		PID_7.a	

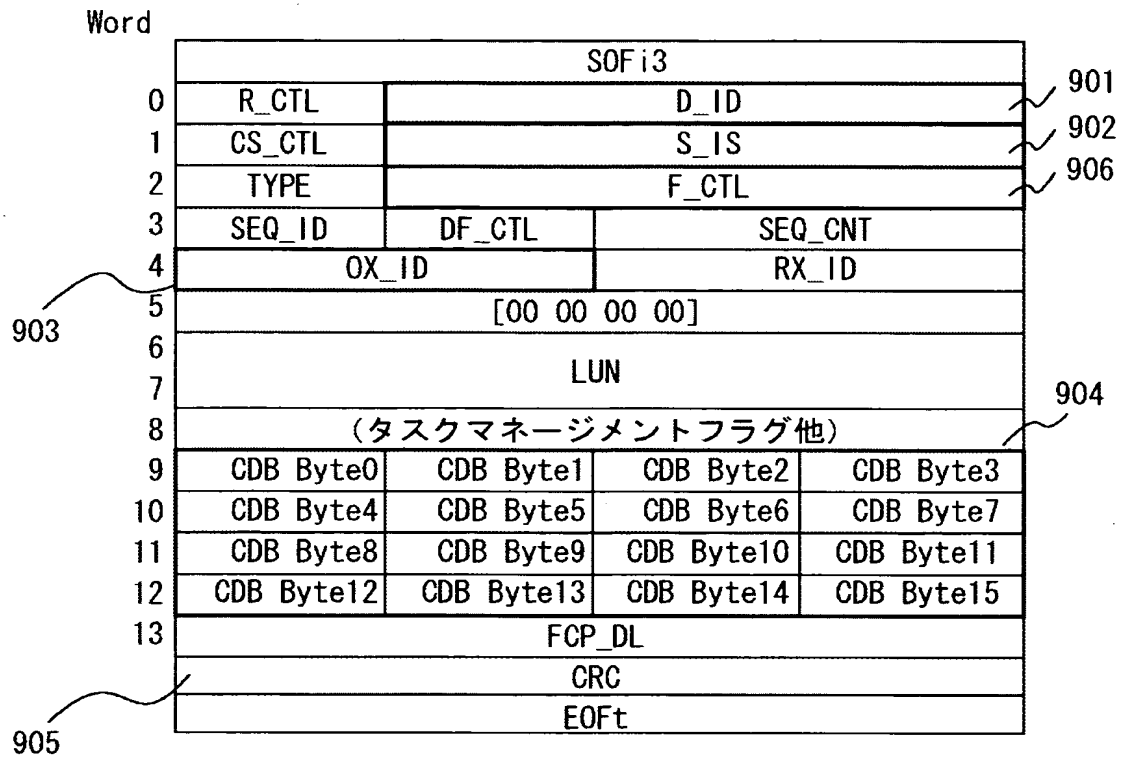
【図 8】

図 8



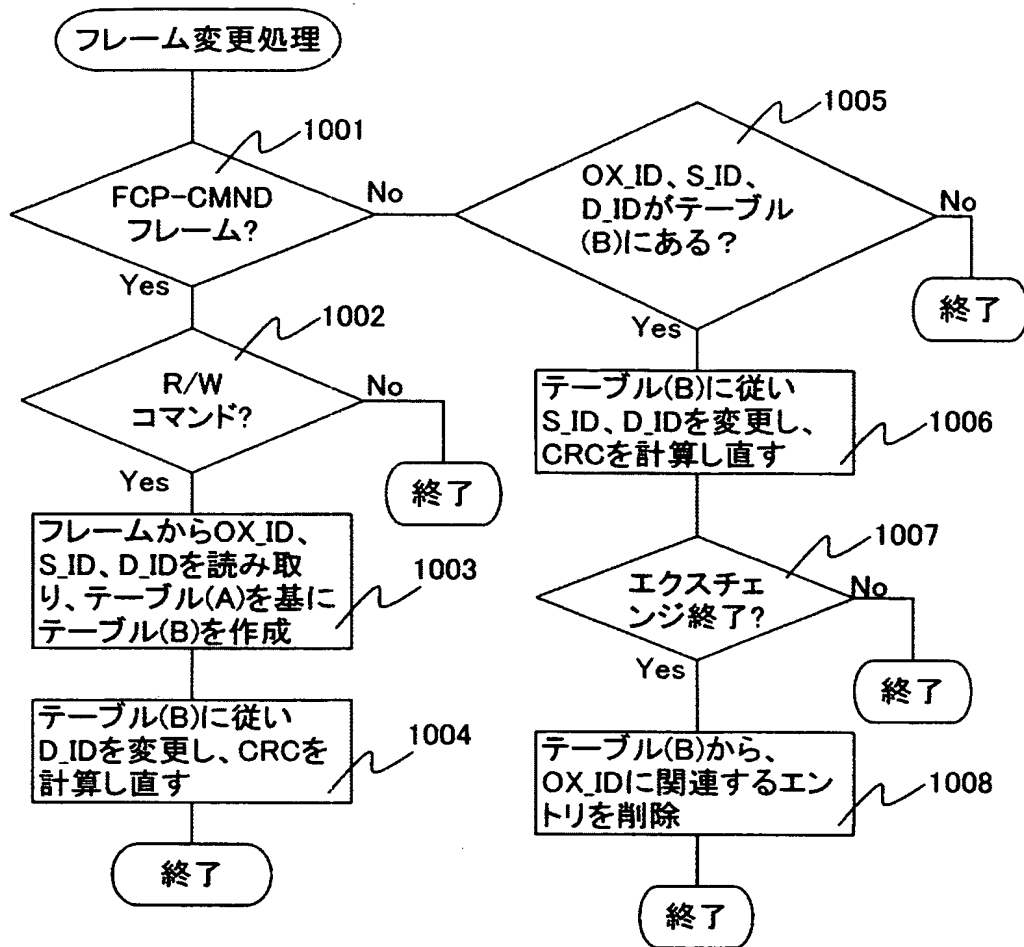
【図 9】

図 9



【図 10】

図10



【図 11 A】

図11A

1101 Drive Port_ID	1102 Read	1103 Write
PID 0.a	PID 0.a	PID 0.b
PID 0.b	PID 0.a	PID 0.b
PID 1.a	PID 1.a	PID 1.b
PID 1.b	PID 1.a	PID 1.b
PID 2.a	PID 2.a	PID 2.b
PID 2.b	PID 2.a	PID 2.b
PID 3.a	PID 3.a	PID 3.b
PID 3.b	PID 3.a	PID 3.b
PID 4.a	PID 4.a	PID 4.b
PID 4.b	PID 4.a	PID 4.b

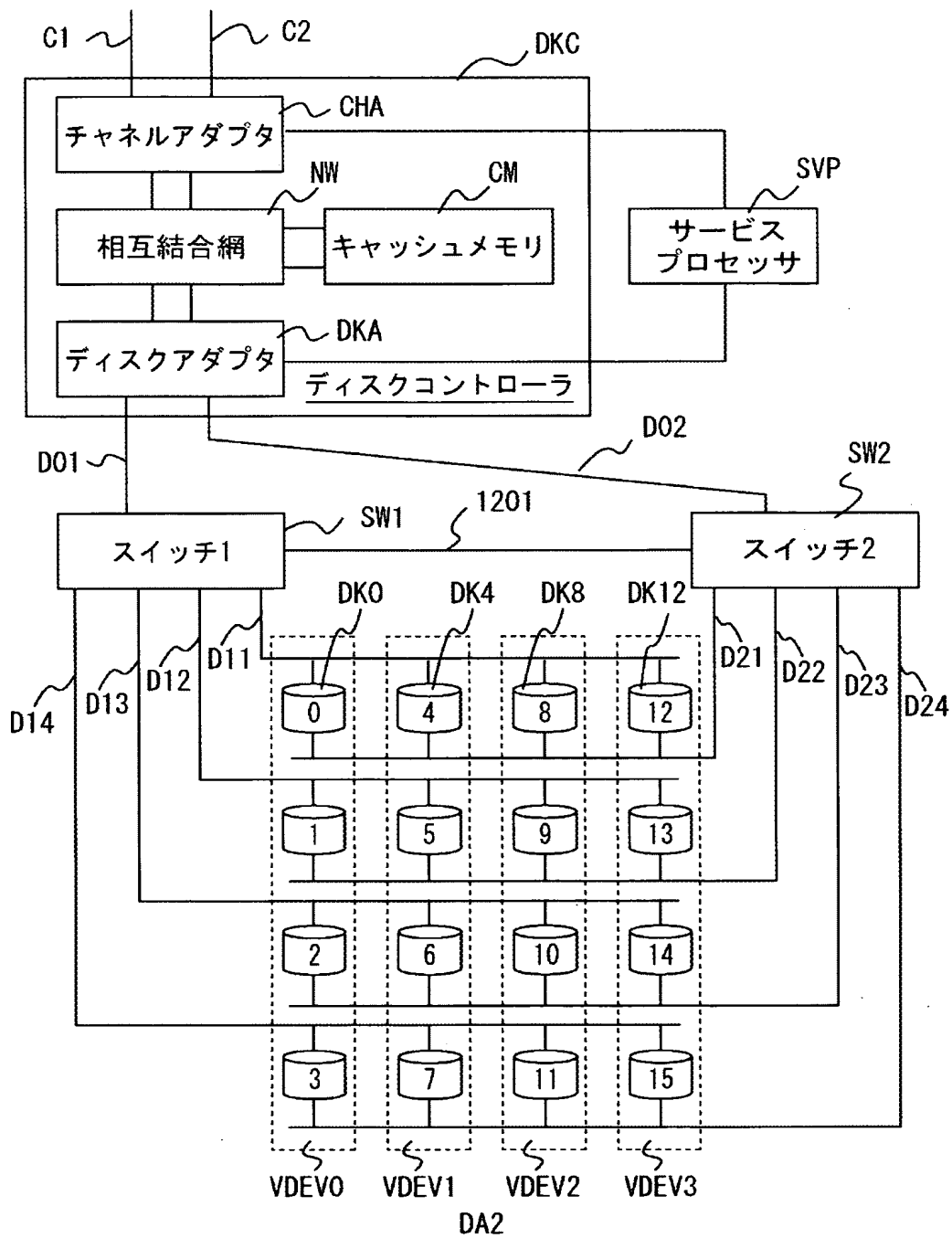
【図 11 B】

図11B

1104 1105 変更前		1106	1107 変更後	1108
OX_ID	S_ID	D_ID	S_ID	D_ID
1109 0x03F2	a	b	変更無	c
1110 0x03F2	c	a	b	無変換
0x03F3	a	e	変更無	f
0x03F3	f	a	e	変更無

【図 12】

図 12



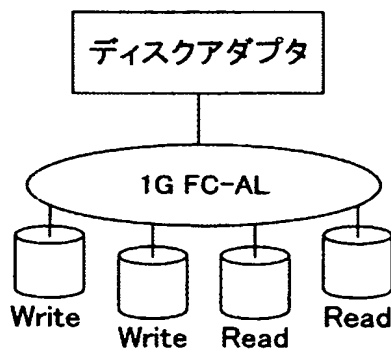
【図 13】

図 13

1301		1302			1303		1304	
		SW1、2正常			SW1障害		SW2障害	
Drive 番号	VDEV	DKA Port	Read Drive Port	Write Drive Port	DKA Port	R/W Drive Port	DKA Port	R/W Drive Port
0	0	0	PID_0.a	PID_0.b	1	PID_0.b	0	PID_0.a
1			PID_1.a	PID_1.b		PID_1.b		PID_1.a
2			PID_2.a	PID_2.b		PID_2.b		PID_2.a
3			PID_3.a	PID_3.b		PID_3.b		PID_3.a
4	1	0	PID_4.a	PID_4.b	1	PID_4.b	0	PID_4.a
5			PID_5.a	PID_5.b		PID_5.b		PID_5.a
6			PID_6.a	PID_6.b		PID_6.b		PID_6.a
7			PID_7.a	PID_7.b		PID_7.b		PID_7.a
8	2	0	PID_8.a	PID_8.b	1	PID_8.b	0	PID_8.a
9			PID_9.a	PID_9.b		PID_9.b		PID_9.a
10			PID_10.a	PID_10.b		PID_10.b		PID_10.a
11			PID_11.a	PID_11.b		PID_11.b		PID_11.a
12	3	0	PID_12.a	PID_12.b	1	PID_12.b	0	PID_12.a
13			PID_13.a	PID_13.b		PID_13.b		PID_13.a
14			PID_14.a	PID_14.b		PID_14.b		PID_14.a
15			PID_15.a	PID_15.b		PID_15.b		PID_15.a

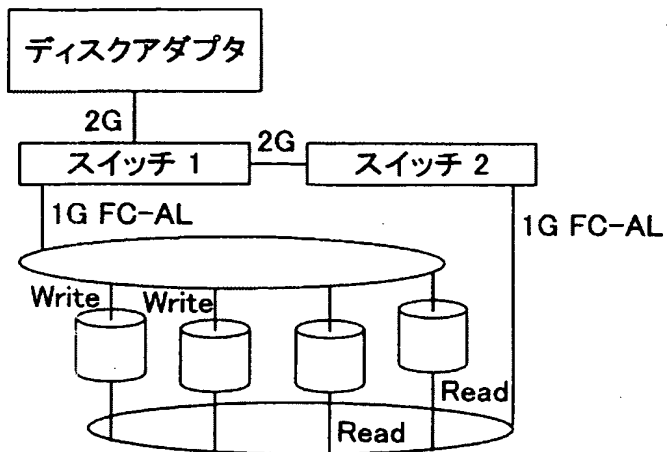
【図 14 A】

図 14A



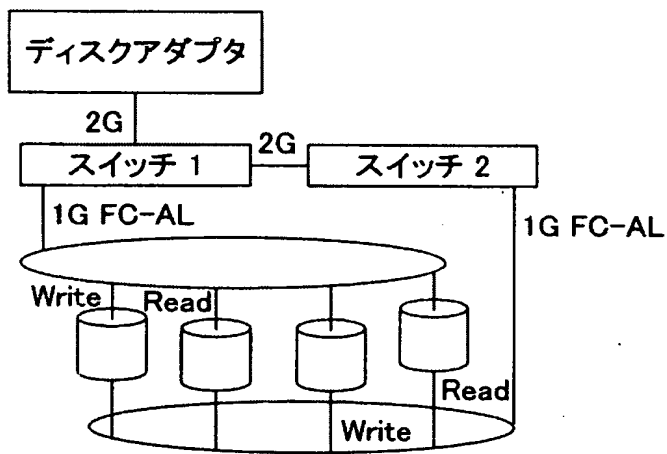
【図 14 B】

図 14B



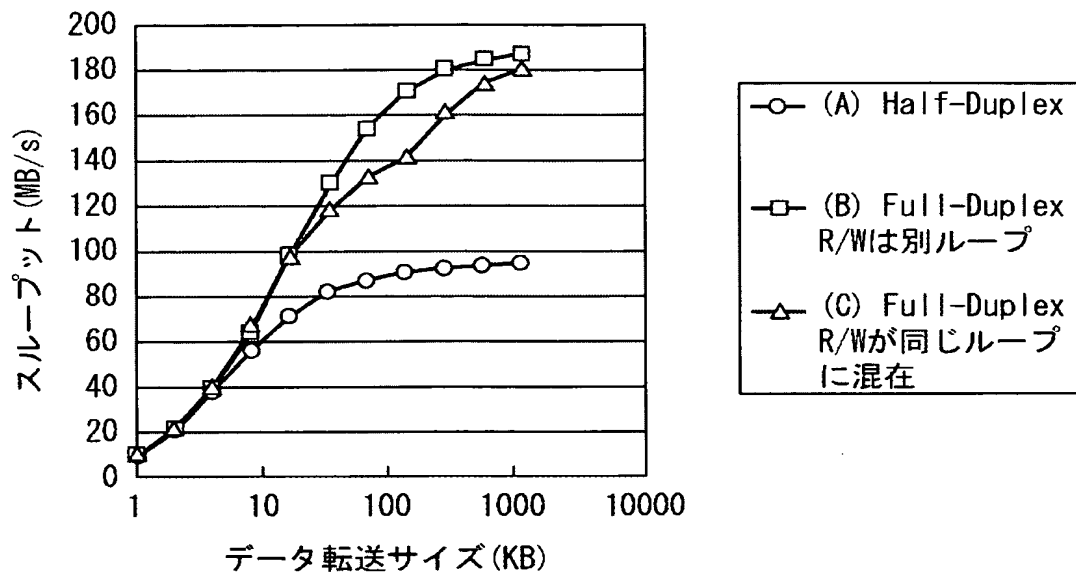
【図 14 C】

図 14C



【図 15】

図 15



【図 16】

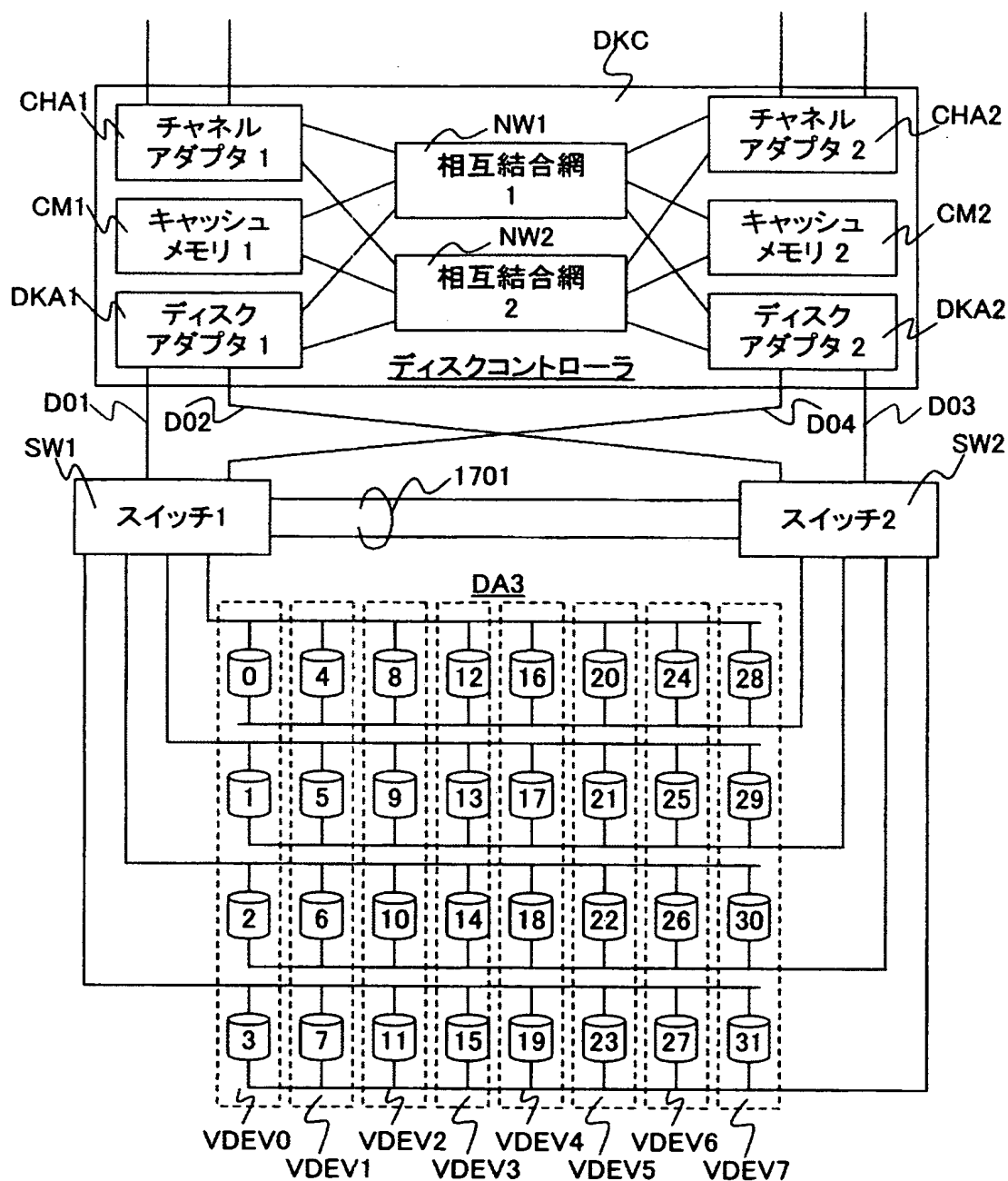
図 16

1601

Drive 番号	VDEV	SW1、2正常			SW1障害		SW2障害	
		DKA Port	Read Drive Port	Write Drive Port	DKA Port	R/W Drive Port	DKA Port	R/W Drive Port
0	0	0	PID_0.a	PID_0.b	1	PID_0.b	0	PID_0.a
1			PID_1.a	PID_1.b		PID_1.b		PID_1.a
2			PID_2.a	PID_2.b		PID_2.b		PID_2.a
3			PID_3.a	PID_3.b		PID_3.b		PID_3.a
4	1	1	PID_4.a	PID_4.b	1	PID_4.b	0	PID_4.a
5			PID_5.a	PID_5.b		PID_5.b		PID_5.a
6			PID_6.a	PID_6.b		PID_6.b		PID_6.a
7			PID_7.a	PID_7.b		PID_7.b		PID_7.a
8	2	0	PID_8.a	PID_8.b	1	PID_8.b	0	PID_8.a
9			PID_9.a	PID_9.b		PID_9.b		PID_9.a
10			PID_10.a	PID_10.b		PID_10.b		PID_10.a
11			PID_11.a	PID_11.b		PID_11.b		PID_11.a
12	3	1	PID_12.a	PID_12.b	1	PID_12.b	0	PID_12.a
13			PID_13.a	PID_13.b		PID_13.b		PID_13.a
14			PID_14.a	PID_14.b		PID_14.b		PID_14.a
15			PID_15.a	PID_15.b		PID_15.b		PID_15.a

【図 17】

図17



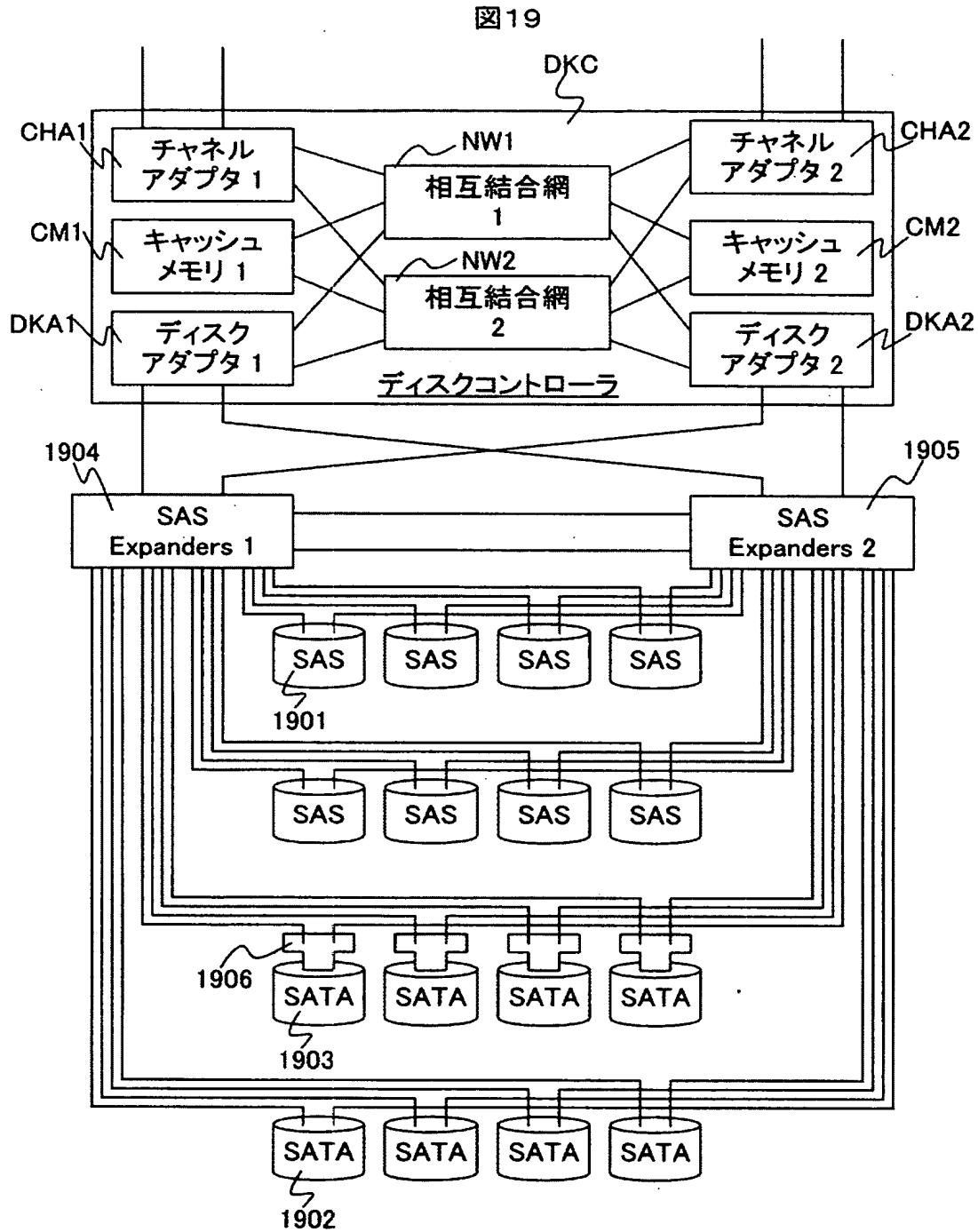
【図 18】

図 18

1801

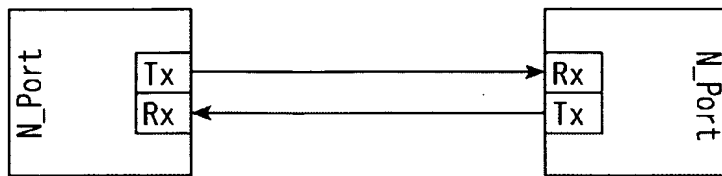
Drive 番号	VDEV	DKA 番号	SW1、2正常			SW1障害		SW2障害	
			DKA Port	Read Drive Port	Write Drive Port	DKA Port	R/W Drive Port	DKA Port	R/W Drive Port
0	0	0	0	PID_0. a	PID_0. b	1	PID_0. b	0	PID_0. a
1				PID_1. a	PID_1. b		PID_1. b		PID_1. a
2				PID_2. a	PID_2. b		PID_2. b		PID_2. a
3				PID_3. a	PID_3. b		PID_3. b		PID_3. a
4	1	1	0	PID_4. a	PID_4. b	0	PID_4. b	1	PID_4. a
5				PID_5. a	PID_5. b		PID_5. b		PID_5. a
6				PID_6. a	PID_6. b		PID_6. b		PID_6. a
7				PID_7. a	PID_7. b		PID_7. b		PID_7. a
8	2	0	1	PID_8. a	PID_8. b	1	PID_8. b	0	PID_8. a
9				PID_9. a	PID_9. b		PID_9. b		PID_9. a
10				PID_10. a	PID_10. b		PID_10. b		PID_10. a
11				PID_11. a	PID_11. b		PID_11. b		PID_11. a
12	3	1	1	PID_12. a	PID_12. b	0	PID_12. b	1	PID_12. a
13				PID_13. a	PID_13. b		PID_13. b		PID_13. a
14				PID_14. a	PID_14. b		PID_14. b		PID_14. a
15				PID_15. a	PID_15. b		PID_15. b		PID_15. a
16	4	0	0	PID_16. a	PID_16. b	1	PID_16. b	0	PID_16. a
17				PID_17. a	PID_17. b		PID_17. b		PID_17. a
18				PID_18. a	PID_18. b		PID_18. b		PID_18. a
19				PID_19. a	PID_19. b		PID_19. b		PID_19. a
20	5	1	0	PID_20. a	PID_20. b	0	PID_20. b	1	PID_20. a
21				PID_21. a	PID_21. b		PID_21. b		PID_21. a
22				PID_22. a	PID_22. b		PID_22. b		PID_22. a
23				PID_23. a	PID_23. b		PID_23. b		PID_23. a
24	6	0	1	PID_24. a	PID_24. b	1	PID_24. b	0	PID_24. a
25				PID_25. a	PID_25. b		PID_25. b		PID_25. a
26				PID_26. a	PID_26. b		PID_26. b		PID_26. a
27				PID_27. a	PID_27. b		PID_27. b		PID_27. a
28	7	1	1	PID_28. a	PID_28. b	0	PID_28. b	1	PID_28. a
29				PID_29. a	PID_29. b		PID_29. b		PID_29. a
30				PID_30. a	PID_30. b		PID_30. b		PID_30. a
31				PID_31. a	PID_31. b		PID_31. b		PID_31. a

【図 19】



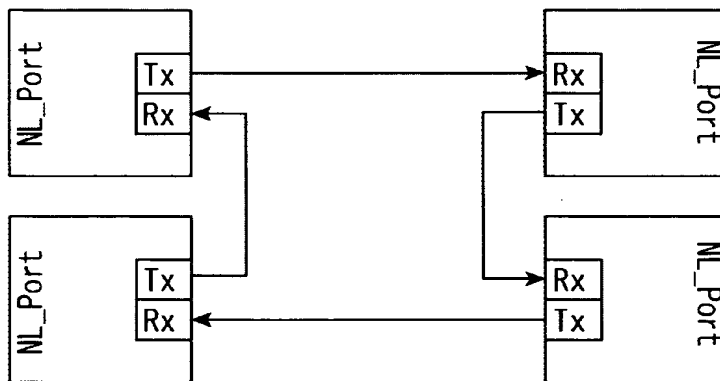
【図 20】

図 20



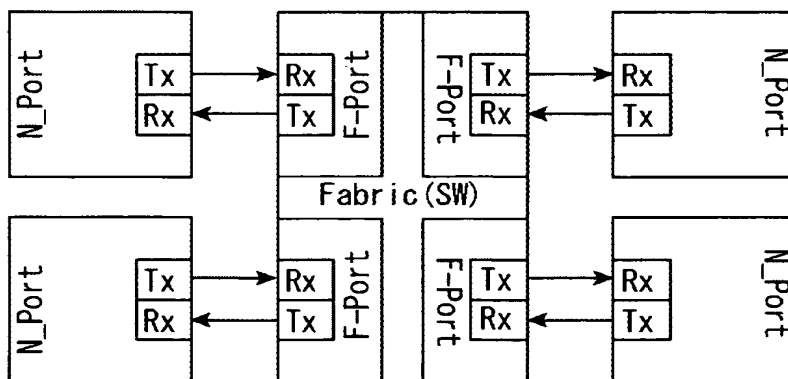
【図 21】

図 21



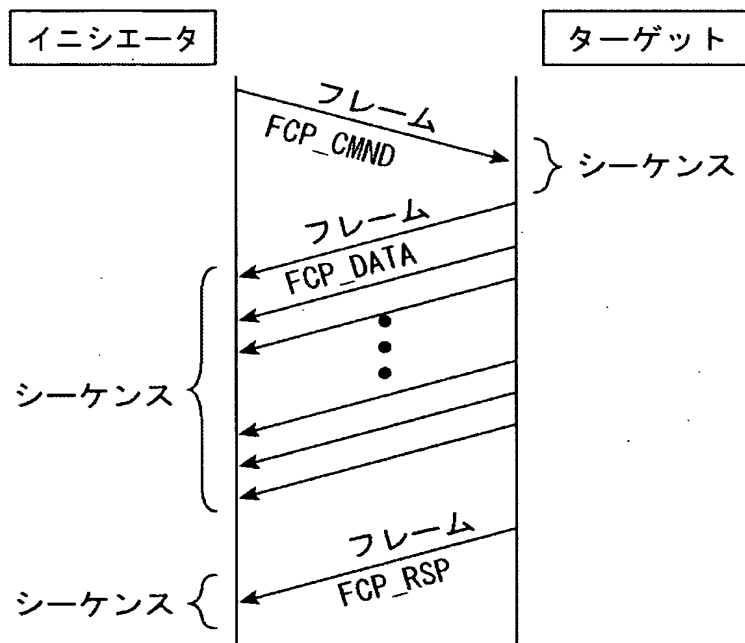
【図 22】

図 22



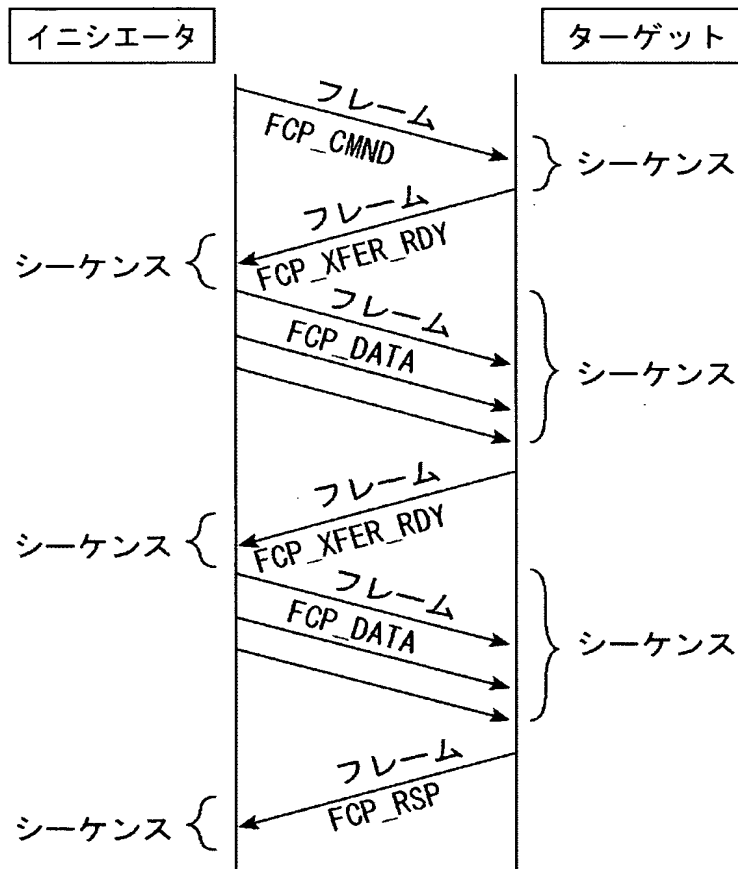
【図 23】

図 23



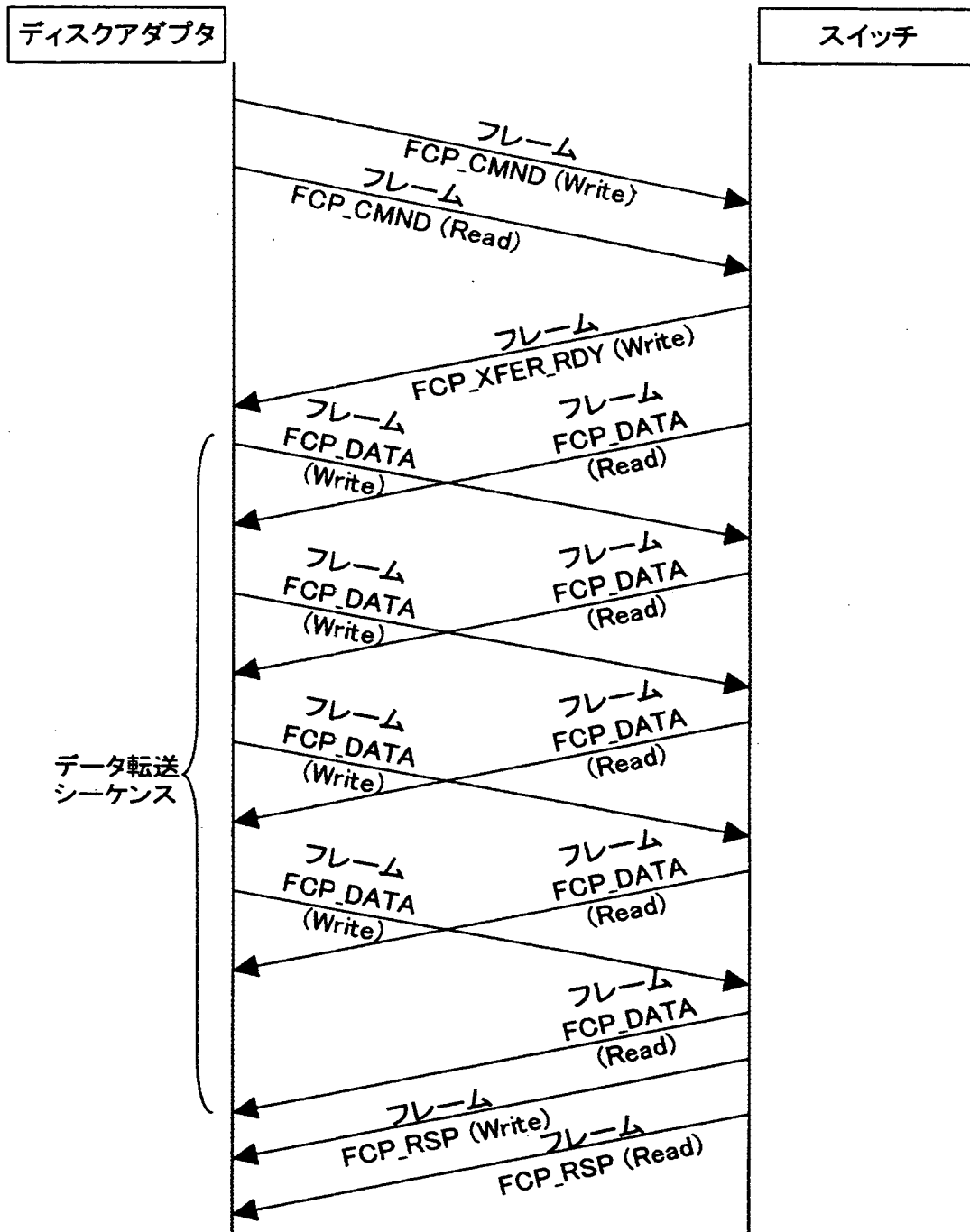
【図 24】

図 24



【図 25】

図25



【図 26】

図 26

Drive 番号	VDEV	DKA 番号	SW1、2正常			SW1障害		SW2障害	
			DKA Port	Read Drive Port	Write Drive Port	DKA Port	R/W Drive Port	DKA Port	R/W Drive Port
0	0	0	0	PID_0.a	PID_0.b	1	PID_0.b	0	PID_0.a
1				PID_1.b	PID_1.a		PID_1.b		PID_1.a
2				PID_2.a	PID_2.b		PID_2.b		PID_2.a
3				PID_3.b	PID_3.a		PID_3.b		PID_3.a
4	1	1	0	PID_4.a	PID_4.b	0	PID_4.b	1	PID_4.a
5				PID_5.b	PID_5.a		PID_5.b		PID_5.a
6				PID_6.a	PID_6.b		PID_6.b		PID_6.a
7				PID_7.b	PID_7.a		PID_7.b		PID_7.a
8	2	0	1	PID_8.a	PID_8.b	1	PID_8.b	0	PID_8.a
9				PID_9.b	PID_9.a		PID_9.b		PID_9.a
10				PID_10.a	PID_10.b		PID_10.b		PID_10.a
11				PID_11.b	PID_11.a		PID_11.b		PID_11.a
12	3	1	1	PID_12.a	PID_12.b	0	PID_12.b	1	PID_12.a
13				PID_13.b	PID_13.a		PID_13.b		PID_13.a
14				PID_14.a	PID_14.b		PID_14.b		PID_14.a
15				PID_15.b	PID_15.a		PID_15.b		PID_15.a
16	4	0	0	PID_16.a	PID_16.b	1	PID_16.b	0	PID_16.a
17				PID_17.b	PID_17.a		PID_17.b		PID_17.a
18				PID_18.a	PID_18.b		PID_18.b		PID_18.a
19				PID_19.b	PID_19.a		PID_19.b		PID_19.a
20	5	1	0	PID_20.a	PID_20.b	0	PID_20.b	1	PID_20.a
21				PID_21.b	PID_21.a		PID_21.b		PID_21.a
22				PID_22.a	PID_22.b		PID_22.b		PID_22.a
23				PID_23.b	PID_23.a		PID_23.b		PID_23.a
24	6	0	1	PID_24.a	PID_24.b	1	PID_24.b	0	PID_24.a
25				PID_25.b	PID_25.a		PID_25.b		PID_25.a
26				PID_26.a	PID_26.b		PID_26.b		PID_26.a
27				PID_27.b	PID_27.a		PID_27.b		PID_27.a
28	7	1	1	PID_28.a	PID_28.b	0	PID_28.b	1	PID_28.a
29				PID_29.b	PID_29.a		PID_29.b		PID_29.a
30				PID_30.a	PID_30.b		PID_30.b		PID_30.a
31				PID_31.b	PID_31.a		PID_31.b		PID_31.a

【書類名】 要約書**【要約】**

【課題】 全二重データ転送に適したバックエンド・ネットワークを有する、ディスク装置およびその制御方法を提供する。

【解決手段】 ディスクアダプタとデュアルポートを有するディスクドライブを、スイッチを介してデュアルループ接続し、ディスクアダプタがディスクドライブへ発行するコマンドの種類（Read/Write）に応じてコマンドの送信先ループを決定し、且つReadのエキステンシブとWriteのエキステンシブが平行して進行するようにディスクアダプタはReadおよびWriteコマンドを発行する。

【選択図】 図 1

特願 2 0 0 3 - 3 8 6 2 8 7

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所